nature

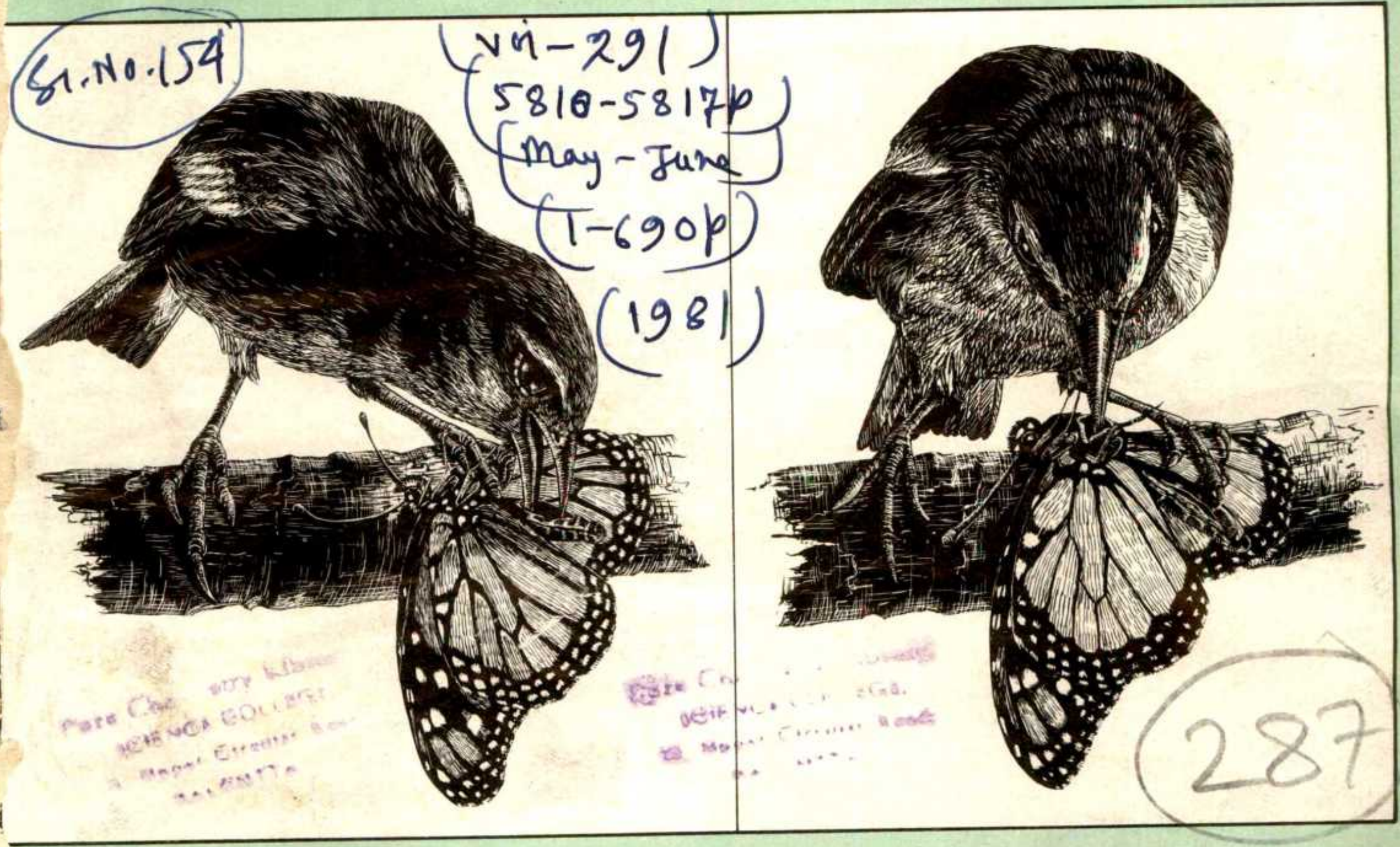
Vol 291 No 5810 7–13 May 1981 £1.25 $3.00

15/1/82

Sl.No.154

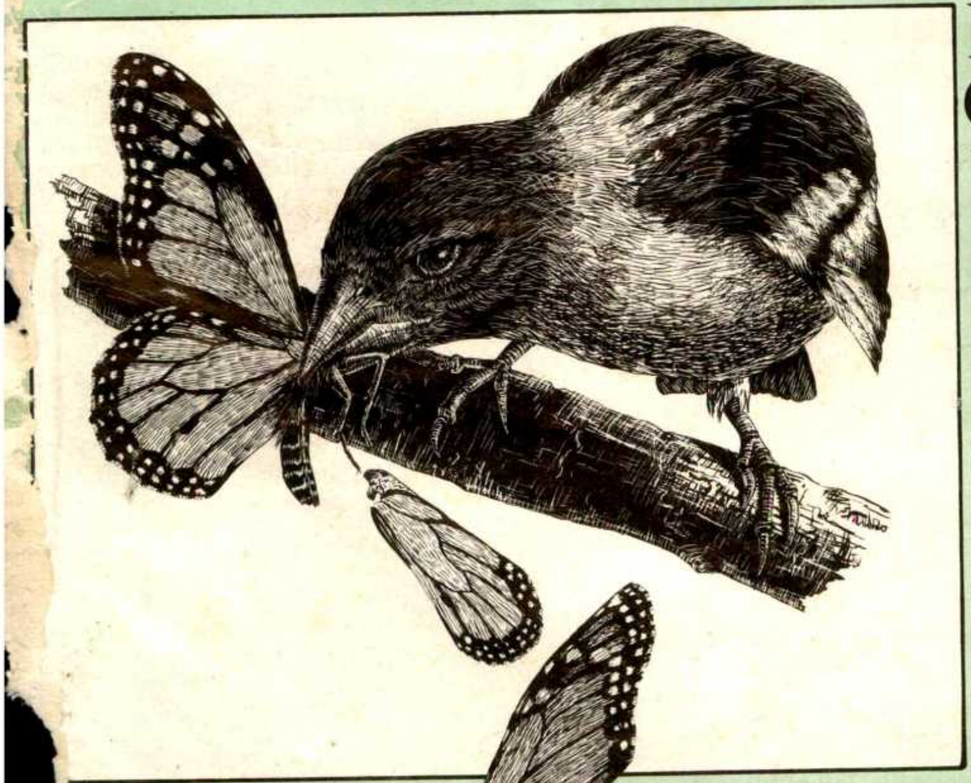
(Vol-291)
(5810-5817p)
(May-June)
(1-690p)
(1981)

287



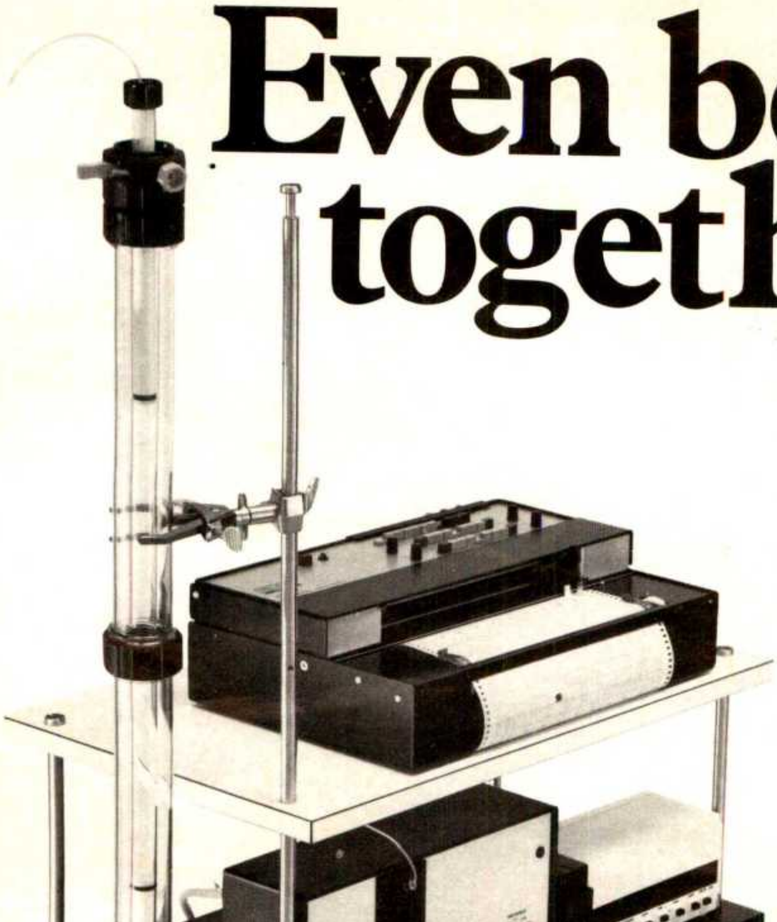
Pure Chemistry Library
SCIENCE COLLEGE
92, Upper Circular Road
CALCUTTA

Pure Ch[illegible]
SCIENCE COLLEGE
92, Upper Circular Road



PREDATION OF UNPALATABLE BUTTERFLIES

Even better together.



Individually, LKB chromatography instruments provide outstanding reliability and consistently accurate results.

But together, they form total, integrated systems that are unequalled in their ease of use and dependability. The LKB 2111 MultiRac® Fraction Collector is the core of just such a system.

LKB excellence in design

The LKB 2111-MultiRac® system combines the ultra-reliable 2111 Fraction Collector with the new MicroPerpex® peristaltic pump, 2210 dual channel recorder and high sensitivity 2138 Uvicord®S UV monitor in one compact package. This entire system takes up less than 2 square feet of bench or cold room space. Yet, every instrument is easy to reach and adjust.

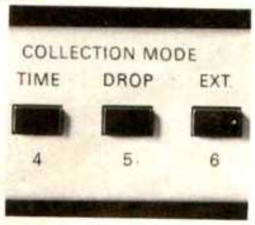
Integrated ease of operation

With the programmable 2111 MultiRac controlling the run, you're assured of accurate, no waste, no problem collections. And with precise pump control of fraction volumes, automatic flow stop function during tube changes and continuous display of flow rate on the MicroPerpex pump, the system needs minimal interaction time with laboratory personnel.

Versatile and reliable

This new LKB collector draws on years of design experience and incorporates the most advanced circuitry. Because the MultiRac precisely shifts the drop head, it can accommodate a large variety of collection vessels.

The sophisticated microprocessor controller of the collector lets you select time, drop or precise volume modes, diverts void volumes to waste and can be programmed to save tubes by collecting valleys in large volumes, peaks in small.



See it all together

Ask your LKB representative for more information on LKB complete chromatography systems. See how they work even better together.

LKB

Head office:
LKB Produkter AB
Box 305
S-161 26 Bromma, Sweden
Tel. 08/98 00 40, telex 10492

UK sales office:
LKB Instruments, Ltd.
232 Addington Road
South Croydon, Surrey CR28YD
England

Other sales offices in:
Athens (for the Middle East),
Copenhagen, Ghent,
The Hague, Hong Kong,
London, Munich, Paris,
Rome, Turku, Vienna

Circle No. 06 on Reader Enquiry Card.

nature

Volume 291 No.5810 7 May 1981

NEWS

Reagan's need of a proliferation policy 1

Advice and dissent on civil science 2

Recombinant DNA guidelines wear thin: NIH moves to voluntary regulation 3

Plea to keep guidelines 3

Role of advisory committee questioned 3

Carter's science team: Yesterday's men 4

Ariane development: Launch at last? 4

British aerospace: Limbering up 4

European Parliament: Nuclear ban debate 5

East Germany banks on research 5

British countryside: Ever onward 5

Reagan's cuts: Handler objects 6

Nuclear power in Yugoslavia: Staying independent 6

Proton decay hunt 6

Correspondence: Conservation sites (A.H. Bretag); Literature search (A. Brooks); Avoiding fraud (J.R. Ravetz); Cladistic clues (W. Kohlberger); Cafeteria feeding (J.R. Baker); Ten years of natural history (T.R.E. Southwood and R.H. Hedley) 7,8

NEWS AND VIEWS

Spatial frequency analysis in vision: Oliver Braddick justifies this approach 9

Sliding charge density waves: P.A. Lee discusses evidence for their existence 11

Mobile genetic elements: new Soviet studies: A.D. Mirzabekov reports a Moscow conference 12

Darwin's missing notebooks come to light: Duncan M. Porter describes the Cambridge discovery 13

The well-managed lawn: Peter D. Moore assesses energy expenditure on this type of grassland 14

The Twelfth Lunar and Planetary Science Conference:

Carbon in carbonaceous chondrites: carbyne or graphite? P.P.K. Smith 15

Impact cratering: Richard A.F. Grieve 16

ARTICLES

E.J. Jankowski and D.J. Drewry **The structure of West Antarctica from geographical studies** 17

C.E. Ford **Parental liquids of the Skaergaard intrusion cumulates** 21

J.F. Allen, J. Bennett, K.E. Steinback and C.J. Arntzen **Chloroplast protein phosphorylation couples plastoquinone redox state to distribution of excitation energy between photosystems** 25

P.J. Gearhart, N.D. Johnson, R. Douglas and L. Hood **IgG antibodies to phosphorylcholine exhibit more diversity than their IgM counterparts** 29

J.E. Coligan, T.J. Kindt, H. Uehara, J. Martinko and S.G. Nathenson **Primary structure of a murine transplantation antigen** 35

R.W. Jones, J.M. Old, R.J. Trent, J.B. Clegg and D.J. Weatherall **Major rearrangement in the human β-globin gene cluster** 39

LETTERS

D.N. Page and M.R. McKee **Eternity matters** 44

P.H. Fowler, R.N.F. Walker, M.R.W. Masheder, R.T. Moses and A. Worley **Ultra-heavy cosmic ray studies with Ariel VI** 45

G.L. Potter, H.W. Ellsaesser, M.C. MacCracken and J.S. Ellis **Albedo change by man: test of climatic effects** 47

D.W. Oscarson, P.M. Huang, C. Defosse and A. Herbillon **Oxidative power of Mn(IV) and Fe(III) oxides with respect to As(III) in terrestrial and aquatic environments** 50

M. Caputo **Earthquake-induced ground acceleration** 51

J.D. Kramers, C.B. Smith, N.P. Lock, R.S. Harmon and F.R. Boyd **Can Kimberlites be generated from an ordinary mantle?** 53

J.G. Cogley **Late Phanerozoic extent of dry land** 56

C.H. Thompson **Podzol chronosequences on coastal dunes of eastern Australia** 59

Cover: Monarch butterflies are eaten in different ways by their two major predators in the Mexican overwintering sites. The black-backed oriole (*Icterus abeillei*) seizes a butterfly by its wings, holds the butterfly underfoot, and uses its narrow and delicate bill to strip out the abdominal contents (top left) and thoracic muscles (top right). By leaving the exoskeleton intact and by taste selectivity, the oriole avoids ingesting substantial amounts of the emetic cardenolides. In contrast, the black-headed grosbeak (*Pheucticus melanocephalus*, bottom), after dismembering a monarch with its heavy bill, may eat the various body parts in their entirety. Drawings by Katherine Doktor Sargent, based on field observations by L.S. Fink and L.P. Brower (see p.67) and photographs by George D. Lepp and L.P. Brower.

Contents continued overleaf

M

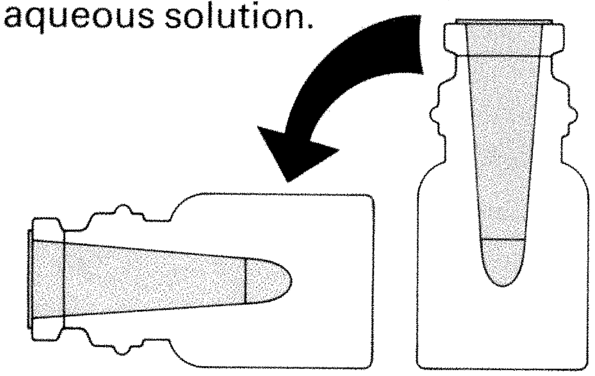
Nature® (ISSN 0028-0836) is published weekly, except the last week in December, by Macmillan Journals Ltd. Annual subscription for USA and Canada US $173. Orders (with remittance) and change of address labels to: Macmillan Journals Ltd, Brunel Rd, Basingstoke RG21 2XS, UK. Second class postage paid at New York, NY 10010 and additional mailing offices. US Postmaster send form 3579 to: *Nature*, 15 East 26 Street, New York, NY10010.

0028-0836/81/19000C-02$01.00

Accidents will happen...

...but all may not be lost!

Amersham's aqueous solutions* of ^{32}P-nucleotides at 10mCi/ml are supplied in "Duoseal V-vials". In an experiment designed to simulate a laboratory accident, 10 open "Duoseal V-vials" were knocked over, each containing 100μl of aqueous solution.



In all cases the surface tension prevented any loss from the specially designed V-insert.

Other advantages of the V-insert

- will accept standard pipette tips
- minimum loss of product, <2% per week lost by absorption
- is not 'wetted' by the solution
- will fit caps from small centrifuge tubes

Aqueous solutions of ^{32}P-nucleotides available weekly from Amersham in "Duoseal V-vials".

Call Amersham or your local representative for further details.

* contains 5mM 2-mercaptoethanol.

~~R516~~
RP417

Amersham

Amersham International Limited, Amersham, England Telephone 024-04-4444
In the U.S.A. and Canada, Amersham Corporation, Illinois 60005 Telephone 312-364-7100 and 800 323-9750 (Tollfree)
In W. Germany, Amersham Buchler GmbH & Co KG, Braunschweig Telephone 05307-4691
Circle No.11 on Reader Enquiry Card.

H.G. Cronblad and B.A. Malmgren **Climatically controlled variation of Sr and Mg in Quaternary planktonic foraminifera** 61

P.D. Taylor **Intra-sex and inter-sex sibling interactions as sex ratio determinants** 64

S.C. Sharma **Retinal projection in a non-visual area after bilateral tectal ablation in goldfish** 66

L.S. Fink and L.P. Brower **Birds can overcome the cardenolide defence of monarch butterflies in Mexico** 67

J.W. Truman, P.H. Taghert, P.F. Copenhaver, N.J. Tublitz and L.M. Schwartz **Eclosion hormone may control all ecdyses in insects** 70

L.S. Honig **Positional signal transmission in the developing chick limb** 72

G.N. Ervin, L.S. Birkemo, C.B. Nemeroff and A.J. Prange Jr **Neurotensin blocks certain amphetamine-induced behaviours** 73

L. Mandarino, D. Stenner, W. Blanchard, S. Nissen, J. Gerich, N. Ling, P. Brazeau, P. Bohlen, F. Esch and R. Guillemin **Selective effects of somatostatin-14, -25 and -28 on *in vitro* insulin and glucagon secretion** 76

W.J. Catalona, T.L. Ratliff and R.E. McCool **γ Interferon induced by *S. aureus* protein A augments natural killing and antibody-dependent CML** 77

F. Marcucci, M. Waller, H. Kirchner and P. Krammer **Production of immune interferon by murine T-cell clones from long-term cultures** 79

D.R. Johnson and M. Jondal **Herpesvirus-transformed cytotoxic T-cell lines** 81

P. Beard, M. Kaneko and P. Cerutti ***N*-acetoxy-acetylaminofluorene reacts preferentially with a control region of intracellular SV40 chromosome** 84

A.A. Rashin **Location of domains in globular proteins** 85

C.S. Craik, S.R. Buchman and S. Beychok **O_2 binding properties of the product of the central exon of β-globin gene** 87

M. Gō **Correlation of DNA exonic regions with protein structural units in haemoglobin** 90

BOOK REVIEWS

Barry Cox **The Making of Mankind (by Richard Leakey) and the TV series of the same title** 93

Vera Rich **The Social Context of Soviet Science (edited by L. L. Lubrano and S.G. Solomon)** 94

P.B. Medawar **The Interaction of Cancer and Host: Its Therapeutic Significance (by Michael F.A. Woodruff)** 95

Richard Lynn **Race, IQ, and Jensen (by James R. Flynn)** 95

W.B. Bonnor **Exact Solutions of Einstein's Field Equations (by D. Kramer *et al.*)** 96

Alan Isaacs **Concise Encyclopedia of the Sciences (edited by John-David Yule)** 96

MISCELLANY

100 years ago 14

New on the market xv

Classified advertising xvii

GUIDE TO AUTHORS

- Review articles should be aimed at a relatively wide readership. Many reviews are invited, but submitted articles may also be accepted; it is advisable to consult us before writing a review article.
- Articles may be up to 3,000 words long with at most six displayed items (figures and tables); they are reports of major research developments.
- Letters are brief reports of original research of unusual and wide interest, not in general longer than 1,000 words; they have at most three or four displayed items.
- 'Matters Arising' permits short discussion (up to 300 words) of papers that have recently appeared in *Nature*.

Articles should be accompanied by an abstract of not more than fifty words. Letters should begin with a paragraph giving the background and main conclusion in terms intelligible to as wide a readership as possible.

Manuscripts may be submitted either to London or New York. Three typed copies should be submitted, each including lettered copies of figures. Typing (including references) should be double spaced. The title should be brief and informative. Pages should be numbered. References, tables and figure legends should start on separate pages. Experimental detail vital to the paper yet which would interrupt the narrative is best placed in the figure legends. Units should be identified in the margin on their first appearance. Equations should occupy single lines if possible: exp (*a*) is preferred to e^a if '*a*' is more than one character.

References are indicated by superscripts in the text. See any contemporary *Nature* for style, but note:

(i) only one reference number need be used if the reference is to several papers by identical authors.

(ii) first and last pages of references should be cited.

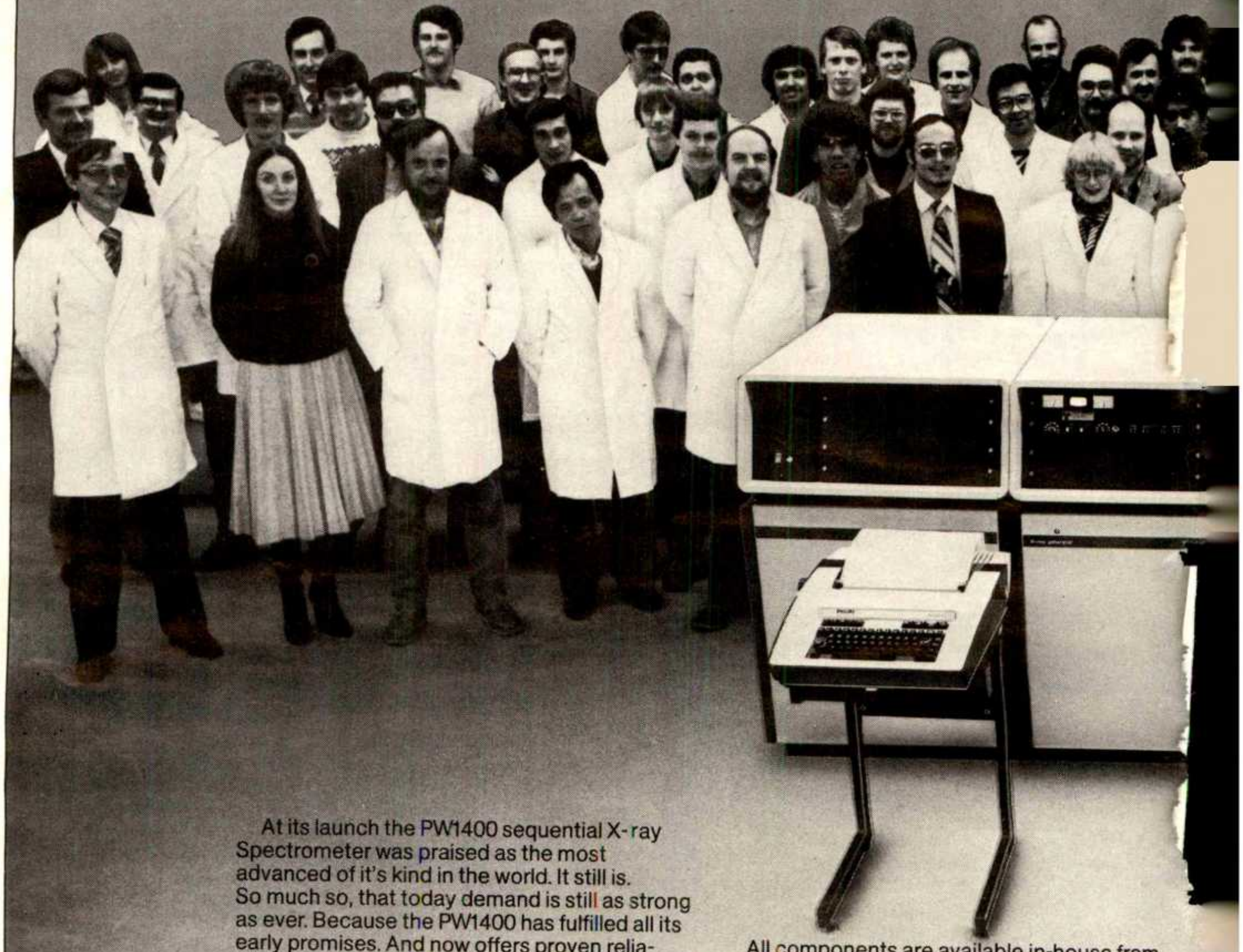
Abbreviations should follow the *World List of Scientific Periodicals*, fourth edn (Butterworth, 1963-65). Symposia are often difficult to refer to and only published or soon-to-be-published volumes should be mentioned in references. Their publisher and place of publication should be clearly indicated 'Personal communication' and 'unpublished' should be incorporated in text.

Artwork should be sent with the manuscript and clearly marked with author's name and the figure number. Line drawings should be either photographic prints or in Indian ink on heavy cartridge paper, tracing paper or similar materials. Most figures are reduced to one column width so originals should be about as wide as a page of *Nature*. To enable figures, particularly maps, to be edited in the same style as the text, they should contain only essential material. ideally, an unlettered original and three lettered copies should be provided; labelling on halftones should, if possible, be avoided entirely. Magnifications quoted should be for the figures as submitted. We are always glad to see artwork for possible use on the cover, but cannot guarantee its return.

In order to save on postal expenses we return only the top copy and artwork of manuscripts that we cannot publish.

Nature's publishing policy is outlined in **258**, 1 (1975) and **264**, v, 11 Nov. (1976).

Why over 250 of the wo
laboratories use one



At its launch the PW1400 sequential X-ray Spectrometer was praised as the most advanced of it's kind in the world. It still is. So much so, that today demand is still as strong as ever. Because the PW1400 has fulfilled all its early promises. And now offers proven reliability. In performance, specification and design, it seems the PW1400 has the happy knack of pleasing all of the people all of the time. It's modular in construction and comes with a choice of generators, sample handlers, data reduction systems and input/output devices.

All components are available in-house from Philips for added reliability. And the system can be connected to a wide range of computers. In terms of back-up we offer extensive software packages, operator training programmes, highly specialised know-how from several application laboratories, and a sophisticated service network.

PHILIPS

Circle No. 08 on Reader Enquiry Card.

nature

Volume 291 No 5811 14 May 1981

[illegible]

NEWS

Will Mitterand be free to change? 97
Calculating flop 98
Tokenism again 98
Congress will battle against proliferation: Senate split confirmation of Malone 99
Science Adviser appointment revived 99
French presidential election: Good for science? 100
French oceanography: Pulling together 100
Bazan protests 100
UK university finance: Truth looms 100
Hungarian research: Cost of elitism 101
Satellite-linked computers: Network planned 101
Research management: Audits go on 101
Indian government policy: Committed to science 102
Arnold at Christie's 102
Third World development: UN funds falter 103
Mary Lindley 103
Correspondence: Progress in Crete (A Kostikas), Xeroderma list (A D Andrews, J L German III & W Clark-Lambert), Be nice to NATO (W M Irvine), Computer algebra D R Stoutemyer), Confusing script (J A J Gowlett) 104

NEWS AND VIEWS

Interferon: decline and stall: Peter Newmark describes current views on the therapeutic potential of interferon 105
Palaeoclimates north and south: M James Salinger reports a conference 106
Hooked on microtubules: Jeremy Hyams describes a new marker of microtubule polarity 107
Magmatic snap, crackle and pop: D H Tarling assesses a new model for magmatic fractionation 108
A catalogue of unidentified objects· David J Thompson on results from the COS B satellite 109
What are mast cells for? T M Dexter, R W Stoddart and S T A Quazzaz review recent research 110
Leishmaniasis and mouse genetics: F E G Cox discusses the nature of this disease 111
Interstellar grains: models and measurements: David A Williams describes recent observations 112

ARTICLES

O van Breeman and B J Bluck — **Episodic granite plutonism in the Scottish Caledonides** 113
A A Gatenby, J A Castleton and M W Saul — **Expression in *E. coli* of maize and wheat chloroplast genes for large subunit of ribulose bisphosphate carboxylase** 117
M Kehoe, R Sellwood, P Shipley and G Dougan — **Genetic analysis of K88-mediated adhesion of enterotoxigenic *Escherichia coli*** 122
P Hudson, J Haley, M Cronk, J Shine and H Niall — **Molecular cloning and characterization of cDNA sequences coding for rat relaxin** 127

LETTERS

W P Blair and R P Kirshner — **Optical spectrum of the unusual supernova remnant G109.1-1.0** 132
D Lindley — **Cosmological consequences of grand unified theories on density fluctuations** 133
M J Cintala — **Meteoroid impact into short-period comet nuclei** 134
H Shimamura and H Watanabe — **Coseismic changes in groundwater temperature of the Usu volcanic region** 137
S C Porter — **Recent glacier variations and volcanic eruptions** 139
P Wellman and R J Tingey — **Glaciation, erosion and uplift over part of East Antarctica** 142
M P Landy and D A Peel — **Short-term fluctuations in heavy metal concentrations in Antarctic snow** 144
B Harte, P M Jackson and R M Macintyre — **Age of mineral equilibria in granulite facies nodules from kimberlites** 147

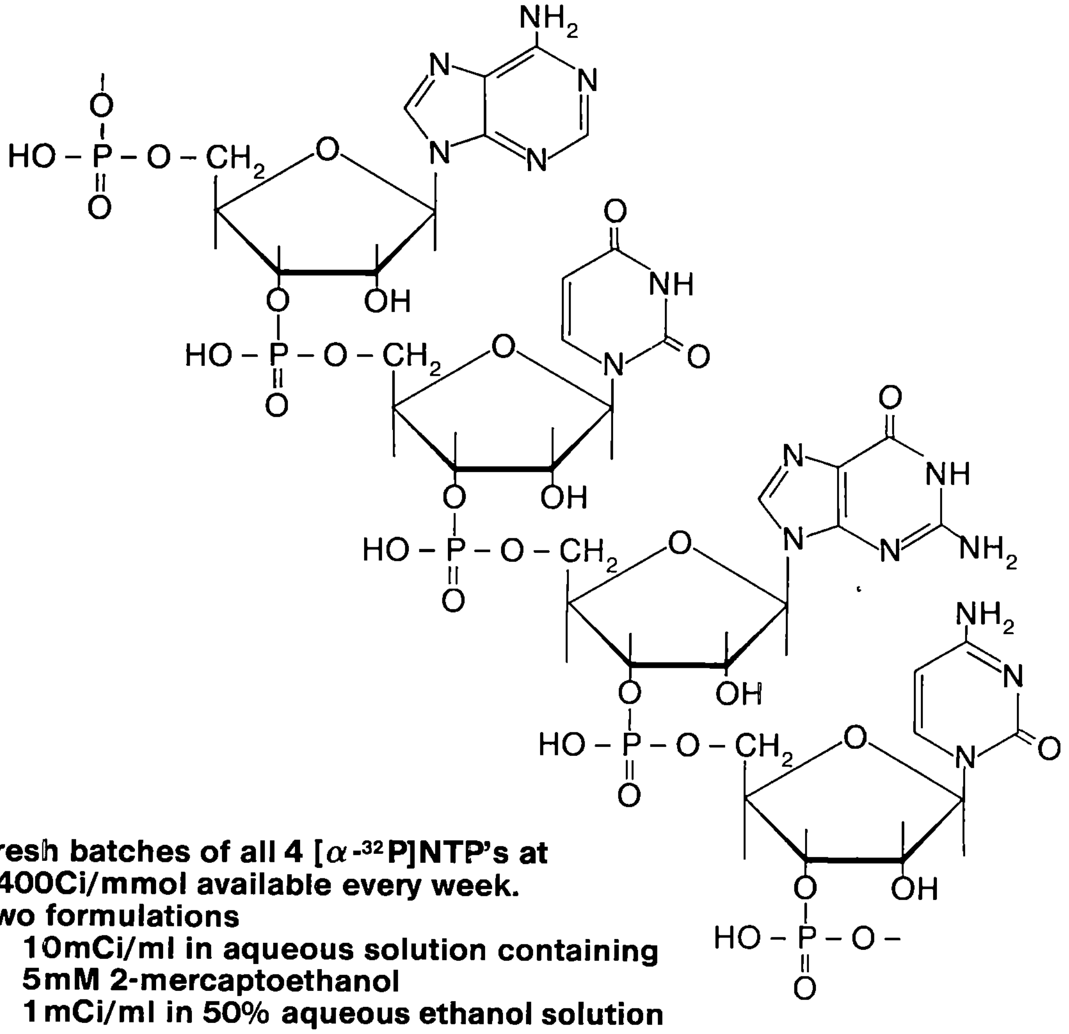
Contents continued overleaf

Cover Collecting snow samples for global pollution studies from the Antarctic ice sheet The results are reported on page 144

Nature© (ISSN 0028-0836) is published weekly, except the last week in December, by Macmillan Journals Ltd Annual subscription for USA and Canada US $173 Orders (with remittance) and change of address labels to Macmillan Journals Ltd, Brunel Rd, Basingstoke RG21 2XS, UK Second class postage paid at New York, NY 10010 and additional mailing offices US Postmaster send form 3579 to *Nature*, 15 East 26 Street New York NY10010

0028-0836/81/20000C-02$01 00

Ribonucleoside $5'$-[α-^{32}P] triphosphates Straight from the bottle every week! From 7th May 1981



Fresh batches of all 4 [α-^{32}P]NTP's at >400Ci/mmol available every week.
Two formulations

- **10mCi/ml in aqueous solution containing 5mM 2-mercaptoethanol**
- **1mCi/ml in 50% aqueous ethanol solution**

Radiochemical Purity

- All batches analysed by HPLC

Aqueous solutions at 10mCi/ml

- Supplied in "Duoseal V-vials"
- Reduced need for evaporation of the solvent
- Avoids the risk of decomposition during evaporation
- Reduced handling minimizes risk of wastage, spillage and contamination

Please ensure you have your copy of our despatch schedule
Call Amersham or your local representative for details.

Amersham International Limited Amersham England Telephone 024 04 4444
In the USA and Canada Amersham Corporation Illinois 60005 Telephone 312 364 7100 and 800/323 9750 (Tollfree)
In W Germany Amersham Buchler GmbH & Co KG, Braunschweig Telephone 05307 4691

Amersham

Circle No 12 on Reader Enquiry Card

S J Risch and F R Rickson — **Mutualism in which ants must be present before plants produce food bodies** 149

S M Rosen, A J Fourcin and B C J Moore — **Voice pitch as an aid to lipreading** 150

J G Mather and R R Baker — **Magnetic sense of direction in woodmice for route-based navigation** 152

H Matsuoka, P J Mulrow, R Franco-Saenz and C H Li — **Stimulation of aldosterone production by β-melanotropin** 155

E A S Al-Dujaili, J Hope, F E Estivariz, P J Lowry and C R W Edwards — **Circulating human pituitary pro-γ-melanotropin enhances the adrenal response to ACTH** 156

Y Kitamura, M Yokoyama, H Matsuda, T Ohno and K J Mori — **Spleen colony-forming cell as common precursor for tissue mast cells and granulocytes** 159

W J Mawby, D J Anstee and M J A Tanner — **Immunochemical evidence for hybrid sialoglycoproteins of human erythrocytes** 161

R Miledi and O D Uchitel — **Properties of postsynaptic channels induced by acetylcholine in different frog muscle fibres** 162

H L Nossel — **Relative proteolysis of the fibrinogen Bβ chain by thrombin and plasmin as a determinant of thrombosis** 165

S S Gill and B D Hammock — **Epoxide hydrolase activity in the mitochondrial fraction of mouse liver** 167

MATTERS ARISING

D I Stirling and H Dalton — **Fortuitous oxidations by methane-utilizing bacteria** 169

I J Higgins, D J Best and R C Hammond — **Reply to Stirling and Dalton** 169

E Volkin and P Benson — **Cycloheximide inhibits DNA synthesis in SV40-transformed cells** 170

F Reyes and J May — **Beaming of Jupiter's decametric radiation** 171

J P Raynaud and M Moguilewsky — **Corticosteroid-binding globulin and corticosterone interaction with progesterone receptors** 171

B D Greenstein and H Al-Khouri — **Reply to Raynaud and Moguilewsky** 172

R Nityananda — **Polar cap models for pulsars** 172

H Ardavan — **Reply to Nityananda** 172

BOOK REVIEWS

John Yudkin — **Medical Aspects of Dietary Fibre: A Report of the Royal College of Physicians. Medical Aspects of Dietary Fiber (edited by G.A. Spiller and R. McPherson Kay)** 173

Jay Greenberg — **Cell Biology: A Comprehensive Treatise, Vols 3 and 4 (edited by L. Goldstein and D.M. Prescott)** 174

Joseph Silk — **Physical Cosmology (edited by R. Balian *et al.*)** 174

Kenneth Mellanby — **The IBP Survey of Conservation Sites (edited by A.R. Clapham)** 175

Michael Mingos — **Transition Metal Clusters (edited by B.F.G. Johnson)** 175

Michael F W Festing — **Animal Suffering: The Science of Animal Welfare (by Marian Stamp Dawkins)** 176

MISCELLANY

100 years ago 110

Errata and corrigendum 168

Classified advertising xviii

GUIDE TO AUTHORS

- Review articles should be aimed at a relatively wide readership Many reviews are invited, but submitted articles may also be accepted, it is advisable to consult us before writing a review article
- Articles may be up to 3,000 words long with at most six displayed items (figures and tables), they are reports of major research developments
- Letters are brief reports of original research of unusual and wide interest, not in general longer than 1,000 words, they have at most three or four displayed items
- 'Matters Arising' permits short discussion (up to 300 words) of papers that have recently appeared in *Nature*

Articles should be accompanied by an abstract of not more than fifty words Letters should begin with a paragraph giving the background and main conclusion in terms intelligible to as wide a readership as possible

Manuscripts may be submitted either to London or New York Three typed copies should be submitted, each including lettered copies of figures Typing (including references) should be double spaced The title should be brief and informative Pages should be numbered References, tables and figure legends should start on separate pages Experimental detail vital to the paper yet which would interrupt the narrative is best placed in the figure legends Units should be identified in the margin on their first appearance Equations should occupy single lines if possible exp (a) is preferred to e^a if 'a' is more than one character

References are indicated by superscripts in the text See any contemporary *Nature* for style, but note

(i) only one reference number need be used if the reference is to several papers by identical authors

(ii) first and last pages of references should be cited

Abbreviations should follow the *World List of Scientific Periodicals*, fourth edn (Butterworth, 1963-65) Symposia are often difficult to refer to and only published or soon-to-be-published volumes should be mentioned in references Their publisher and place of publication should be clearly indicated 'Personal communication' and 'unpublished' should be incorporated in text

Artwork should be sent with the manuscript and clearly marked with author's name and the figure number Line drawings should be either photographic prints or in Indian ink on heavy cartridge paper, tracing paper or similar materials Most figures are reduced to one column width so originals should be about as wide as a page of *Nature* To enable figures, particularly maps, to be edited in the same style as the text, they should contain only essential material ideally, an unlettered original and three lettered copies should be provided, labelling on halftones should, if possible, be avoided entirely Magnifications quoted should be for the figures as submitted We are always glad to see artwork for possible use on the cover, but cannot guarantee its return

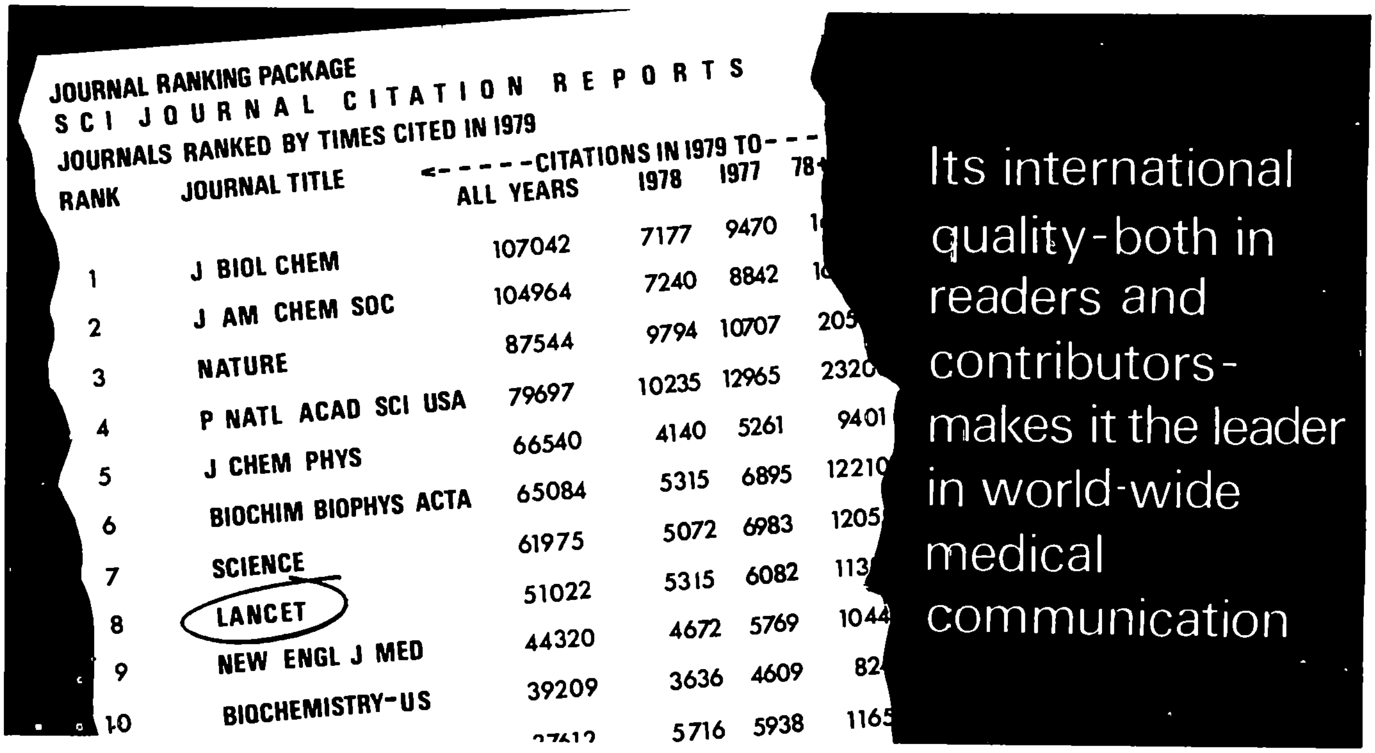
In order to save on postal expenses we return only the top copy and artwork of manuscripts that we cannot publish

Nature's publishing policy is outlined in **258**, 1 (1975) and **264**, v, 11 Nov (1976)

THE LANCET

On the frontiers of medical and scientific knowledge THE LANCET is not only the last word-it's the first!

Consistently ranked by the Institute of Scientific Information as the most cited medical journal in the world.



JOURNAL RANKING PACKAGE
SCI JOURNAL CITATION REPORTS
JOURNALS RANKED BY TIMES CITED IN 1979

RANK	JOURNAL TITLE	ALL YEARS	1978	1977	78
1	J BIOL CHEM	107042	7177	9470	1
2	J AM CHEM SOC	104964	7240	8842	1
3	NATURE	87544	9794	10707	205
4	P NATL ACAD SCI USA	79697	10235	12965	2320
5	J CHEM PHYS	66540	4140	5261	9401
6	BIOCHIM BIOPHYS ACTA	65084	5315	6895	12210
7	SCIENCE	61975	5072	6983	1205
8	LANCET	51022	5315	6082	113
9	NEW ENGL J MED	44320	4672	5769	1044
10	BIOCHEMISTRY-US	39209	3636	4609	82
		[illegible]	5716	5938	1165

Its international quality-both in readers and contributors-makes it the leader in world-wide medical communication

ORDER FORM

To – THE LANCET LTD., 7 Adam Street, Adelphi, London WC2N 6AD.
Please send me THE LANCET each week for one year.

☐ £27.00 in UK and EIRE ☐ Cheque herewith
☐ £40.00 Overseas (Airspeeded) ☐ Invoice me
☐ £15.00 special rate*, (Available in UK and Eire only) for doctors within 5 years of registration or graduation and for medical students.

*Date of registration or student at – ____________________

____

BLOCK CAPITALS PLEASE
NAME ____________________
ADDRESS ____________________
____
SIGNATURE ____________________ DATE ____________________

nature

Volume 291 No 5812 21 May 1981

NEWS

British universities must win lost friends 177
Creationism again an issue in Arkansas: Civil liberties take suit against law 179
Arkansas legislation reflects old quarrels 179
Scopes again? 180
Sakharov conference: Birthday honour 180
UK industrial research: All change 181
Bioengineering hazards: Europe doubts 181
US defence budget: Laser destruction 181
Soviet space: Romanian resources 182
Viewdata systems: Battle joined 182
Coal liquefaction: Trial by post 183
From rank weeds to riches 183
The responsibility of scientists 184
Correspondence: Creationists persist (T H Jukes), Theory upon theory (B Pearson), Use the library (G Croft, I R Winship) 186

NEWS AND VIEWS

Genes of lymphocytes II: Miranda Robertson on T-cell receptors, idiotypes and the major histocompatibility complex 187
Blood lead concentrations reconsidered: Philippe Grandjean on the need for lowering the level of lead exposure 188
Continental crust compromise: Keith Bell assesses new isotopic data from Colorado 189
Clustering of ion channels at the node of Ranvier: L C Fritz and J P Brockes discuss available evidence 190
Magnetic reconnection in the laboratory: Stanley W H Cowley describes a new experiment 191

ARTICLES

D J DePaolo **Neodymium isotopes in the Colorado Front Range and crust–mantle evolution in the Proterozoic** 193
J J Walsh, G T Rowe, R L Iverson and C P McRoy **Biological export of shelf carbon is a sink of the global CO_2 cycle** 196
S W Law and A Dugaiczyk **Homology between the primary structure of α-fetoprotein, deduced from a complete cDNA sequence, and serum albumin** 201
M Salim and B E H Maden **Nucleotide sequence of *Xenopus laevis* 18S ribosomal RNA inferred from gene sequence** 205

LETTERS

G Gilmore and I N Reid **RG0044–2958: a peculiar M supergiant at a distance of 2.5 Mpc** 208
A O Benz, E Fürst, W Hirth and M R Perrenoud **Solar radio blips and X-ray kernels** 210
G Henschen and F Arnold **New positive ion species in the stratosphere** 211
T M L Wigley and P Brimblecombe **Carbon dioxide, ammonia and the origin of life** 213
G S Hoye **Archaeomagnetic secular variation record of Mount Vesuvius** 216
J L Orpen and J F Wilson **Stromatolites at $\sim$3,500 Myr and a greenstone–granite unconformity in the Zimbabwean Archaean** 218
T Elliott and K O Ladipo **Syn-sedimentary gravity slides (growth faults) in the Coal Measures of South Wales** 220
A S Monin, E A Plakhin, A M Podrazhansky, A M Sagalevich and O G Sorokhtin **Visual observations of the Red Sea hot brines** 222
D A Dunn, T C Moore Jr and L D Keigwin Jr **Atlantic-type carbonate stratigraphy in the late Miocene Pacific** 225
K Reisinger, M Stoeppler and H W Nürnberg **Evidence for the absence of biological methylation of lead in the environment** 228
R A Kosher and M P Savage **Glycosaminoglycan synthesis by the apical ectodermal ridge of chick limb bud** 231

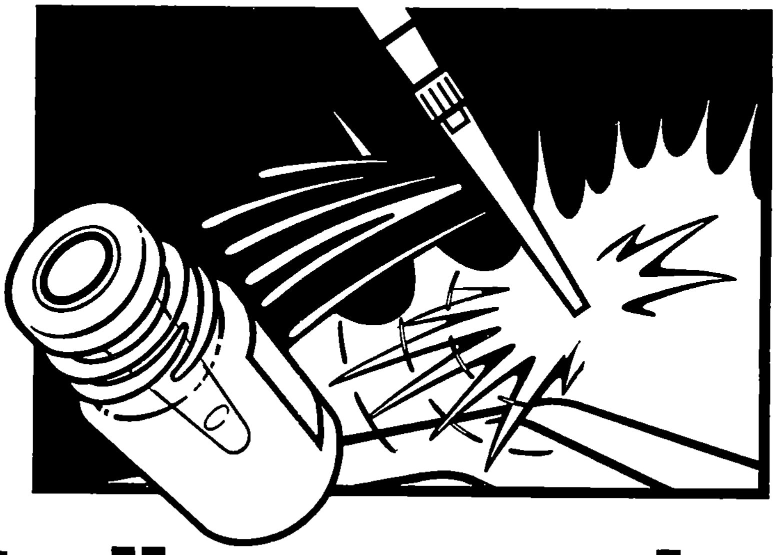
Contents continued overleaf

Cover: Landsat image of an area of the Colorado Front Range west of Denver Isotopic age determinates for rock samples from this area and their implications for crustal evolution are presented on p 193 Photograph courtesy of NASA Goddard Institute for Space Studies, New York

M

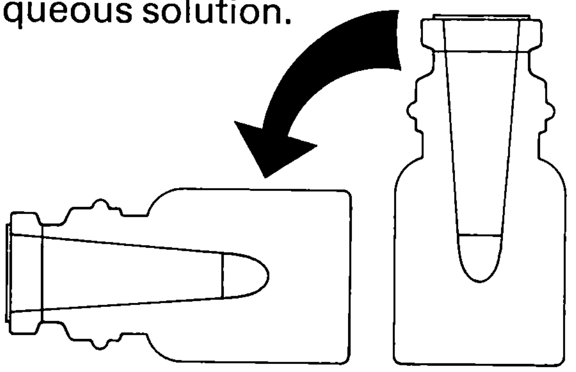
Nature® (ISSN 0028-0836) is published weekly, except the last week in December, by Macmillan Journals Ltd Annual subscription for USA and Canada US $173 Orders (with remittance) and change of address labels to Macmillan Journals Ltd, Brunel Rd, Basingstoke RG21 2XS, UK Second class postage paid at New York, NY 10010 and additional mailing offices US Postmaster send form 3579 to *Nature*, 15 East 26 Street, New York, NY10010

Accidents will happen...



...but all may not be lost!

Amersham's aqueous solutions* of ^{32}P-nucleotides at 10mCi/ml are supplied in "Duoseal V-vials". In an experiment designed to simulate a laboratory accident, 10 open "Duoseal V-vials" were knocked over, each containing 100μl of aqueous solution.



In all cases the surface tension prevented any loss from the specially designed V-insert.

Other advantages of the V-insert

- will accept standard pipette tips
- minimum loss of product, <2% per week lost by absorption
- is not 'wetted' by the solution
- will fit caps from small centrifuge tubes

Aqueous solutions of ^{32}P-nucleotides available weekly from Amersham in "Duoseal V-vials".

Call Amersham or your local representative for further details.

* contains 5mM 2-mercaptoethanol

Amersham International Limited, Amersham England Telephone 024-04-4444
In the U S A and Canada, Amersham Corporation, Illinois 60005 Telephone 312 364 7100 and 800 323 9750 (Tollfree)
In W Germany, Amersham Buchler GmbH & Co KG, Braunschweig Telephone 05307-4691

Circle No 20 on Reader Enquiry Card

Amersham

J A Bluestone, S O Sharrow, S L Epstein, K Ozato and D H Sachs **Induction of anti-H-2 antibodies without alloantigen exposure by *in vivo* administration of anti-idiotype** 233

B L M Hogan, A Taylor and E Adamson **Cell interactions modulate embryonal carcinoma cell differentiation into parietal or visceral endoderm** 235

S K Hoiseth and B A D Stocker **Aromatic-dependent *Salmonella* typhimurium are non-virulent and effective as live vaccines** 238

T A Levitt and M M Salpeter **Denervated endplates have a dual population of junctional acetylcholine receptors** 239

N Franceschini, R Hardie, W Ribi and K Kirschfeld **Sexual dimorphism in a photoreceptor** 241

I Olsen, M F Dean, G Harris and H Muir **Direct transfer of a lysosomal enzyme from lymphoid cells to deficient fibroblasts** 244

L C Andersson, V -P Lehto, S Stenman, R A Badley and I Virtanen **Diazepam induces mitotic arrest at prometaphase by inhibiting centriolar separation** 247

M W Klymkowsky **Intermediate filaments in 3T3 cells collapse after intracellular injection of a monoclonal anti-intermediate filament antibody** 249

D R Hathaway, C R Eaton and R S Adelstein **Regulation of human platelet myosin light chain kinase by the catalytic subunit of cyclic AMP-dependent protein kinase** 252

M Gould-Somero **Localized gating of egg Na^+ channels by sperm** 254

A Spadaro, A Spena, V Santonastaso and P Donini **Stringency without ppGpp accumulation** 256

P L Paine, T W Pearson, L J M Tluczek and S B Horowitz **Nuclear sodium and potassium** 258

H Jornvall, A Carlstrom, T Pettersson, B Jacobsson, M Persson and V Mutt **Structural homologies between prealbumin, gastrointestinal prohormones and other proteins** 261

R E Stenkamp, L C Sieker, L H Jensen and J Sanders-Loehr **Structure of the binuclear iron complex in metazidohaemerythrin from *Themiste dyscritum* at 2.2 Å resolution** 263

MATTERS ARISING

J V Smith and J J Pluth **Si, Al ordering in Linde A zeolite** 265

L A Bursill, E A Lodge and J M Thomas **Reply to Smith and Pluth** 265

B M F Galdikas **Modern adaptations in orang-utans?** 266

R J Smith and D R Pilbeam **Reply to Galdikas** 266

BOOK REVIEWS

Edmund Leach **Genes, Mind, and Culture: The Coevolutionary Process (by C.J. Lumsden and E.O. Wilson)** 267

Roger G Burns **Underwater Minerals (by D.S. Cronan)** 268

M P Atherton **Metamorphic Petrology, 2nd Edn (by F.J. Turner)** 269

John Currey **Skeletal Growth of Aquatic Organisms (edited by D.C. Rhoads and R.A. Lutz)** 269

Robert Fox **The Ferment of Knowledge: Studies in the Historiography of Eighteenth-century Science (edited by G.S. Rousseau and Roy Porter)** 270

MISCELLANY

Author index, Vol 290 xv

Books received xx

Classified advertising xxxv

PRODUCT REVIEW This week's product review (page xxix) covers tissue culture

GUIDE TO AUTHORS

- Review articles should be aimed at a relatively wide readership Many reviews are invited, but submitted articles may also be accepted, it is advisable to consult us before writing a review article
- Articles may be up to 3,000 words long with at most six displayed items (figures and tables), they are reports of major research developments
- Letters are brief reports of original research of unusual and wide interest, not in general longer than 1,000 words, they have at most three or four displayed items
- 'Matters Arising' permits short discussion (up to 300 words) of papers that have recently appeared in *Nature*

Articles should be accompanied by an abstract of not more than fifty words Letters should begin with a paragraph giving the background and main conclusion in terms intelligible to as wide a readership as possible

Manuscripts may be submitted either to London or New York Three typed copies should be submitted, each including lettered copies of figures Typing (including references) should be double spaced The title should be brief and informative Pages should be numbered References, tables and figure legends should start on separate pages Experimental detail vital to the paper yet which would interrupt the narrative is best placed in the figure legends Units should be identified in the margin on their first appearance Equations should occupy single lines if possible exp (a) is preferred to e^a if 'a' is more than one character

References are indicated by superscripts in the text See any contemporary *Nature* for style, but note

(i) only one reference number need be used if the reference is to several papers by identical authors

(ii) first and last pages of references should be cited

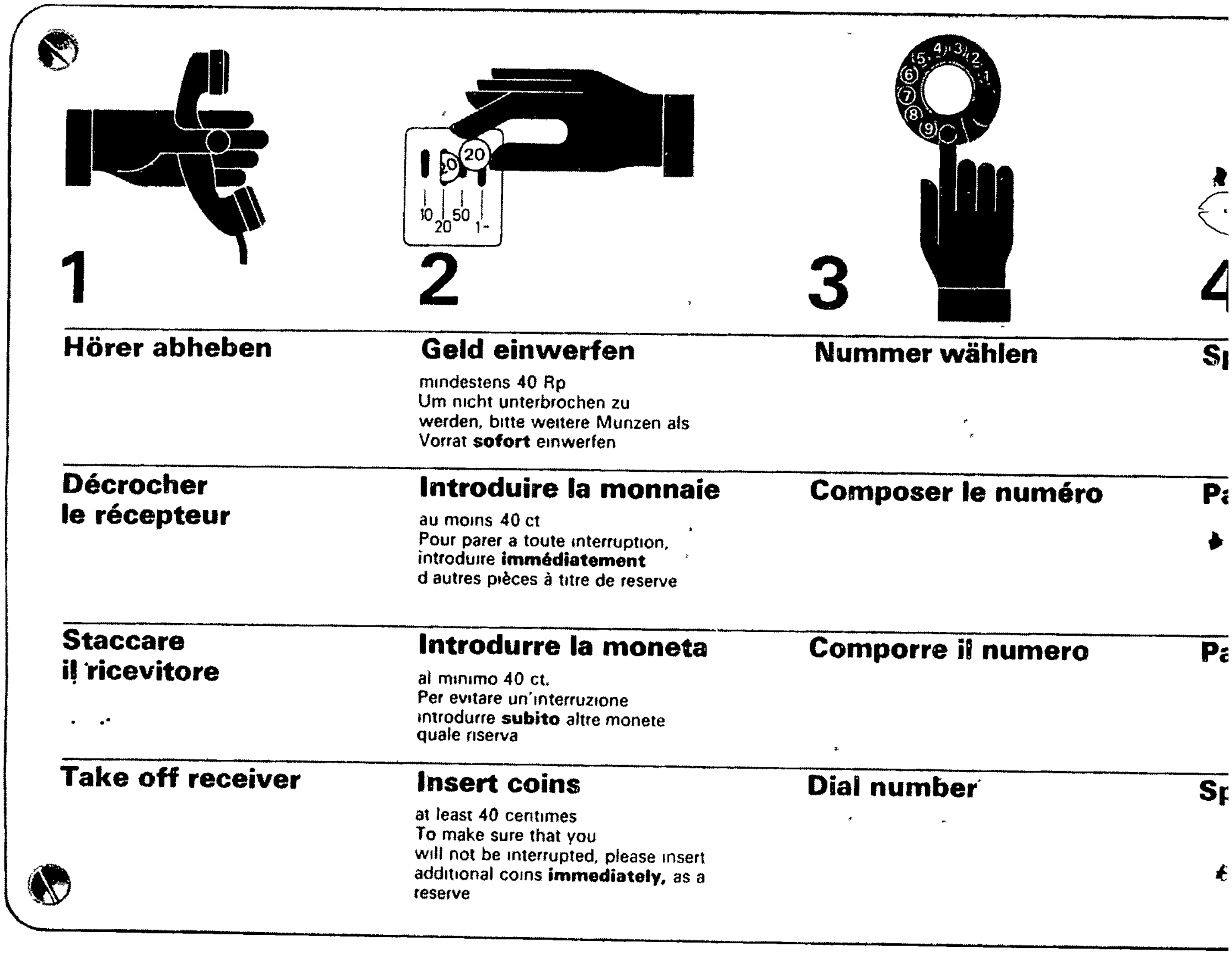
Abbreviations should follow the *World List of Scientific Periodicals*, fourth edn (Butterworth, 1963-65) Symposia are often difficult to refer to and only published or soon-to-be-published volumes should be mentioned in references Their publisher and place of publication should be clearly indicated 'Personal communication' and 'unpublished' should be incorporated in text

Artwork should be sent with the manuscript and clearly marked with author's name and the figure number Line drawings should be either photographic prints or in Indian ink on heavy cartridge paper, tracing paper or similar materials Most figures are reduced to one column width so originals should be about as wide as a page of *Nature* To enable figures, particularly maps, to be edited in the same style as the text, they should contain only essential material ideally, an unlettered original and three lettered copies should be provided, labelling on halftones should, if possible, be avoided entirely Magnifications quoted should be for the figures as submitted We are always glad to see artwork for possible use on the cover but cannot guarantee its return

In order to save on postal expenses we return only the top copy and artwork of manuscripts that we cannot publish

Nature's publishing policy is outlined in **258**, 1 (1975) and **264**, v, 11 Nov (1976)

Tel. 027/831999. Public te
Canton Valais, Switzerl



1	2	3	4
Hörer abheben	**Geld einwerfen** mindestens 40 Rp Um nicht unterbrochen zu werden, bitte weitere Münzen als Vorrat **sofort** einwerfen	**Nummer wählen**	**Sp**
Décrocher le récepteur	**Introduire la monnaie** au moins 40 ct Pour parer à toute interruption, introduire **immédiatement** d'autres pièces à titre de réserve	**Composer le numéro**	**Pa**
Staccare il ricevitore	**Introdurre la moneta** al minimo 40 ct. Per evitare un'interruzione introdurre **subito** altre monete quale riserva	**Comporre il numero**	**Pa**
Take off receiver	**Insert coins** at least 40 centimes To make sure that you will not be interrupted, please insert additional coins **immediately**, as a reserve	**Dial number**	**Sp**

Arolla is at the very back of the Val d'Hérens, separated from Zermatt and the celebrated Matterhorn by 25 km as the crow flies, and by the 4357-m Dent Blanche

Like almost every Swiss hamlet it has a public telephone, of course, and there is nothing unusual about the one at Arolla

You dial direct from here to almost anywhere in Europe (if you have enough change) The telephone functions day and night And if someone should not know how it works, there's a set of instructions

The instructions are in German, French, Italian, and English And this is where things start to get very Swiss indeed

Because our country is in fact quadrilingual, and since we also have to make an effort to understand and to be understood even within this small area, we can also appreciate the position of a guest who speaks a different language Accordingly, as most Swiss are conditioned to expect that strangers will be coming, and they do not want them to feel like strangers, it is no surprise to find the telephone instructions include a version in the world-wide English language, rather than our own fourth national tongue, Rhaeto Romanic

Neither is it a surprise that we have one of the world's most closely meshed rail networks, to help visitors to reach their holiday destinations as fast and

15\12\82

nature

Volume 291 No 5813 28 May 1981

Does creation deserve equal time? 271
Organizing change 272
Too much TV? 272

NEWS 273

Budget cuts cast shadows overseas, Anxieties of Oslo secrets trial, UK nuclear power, Gloom on research, US science funding, High energy physics, Hormone legislation, Canadian radioastronomy, Hungarian agriculture, Direct satellite broadcasting, Cooling by sackcloth

CORRESPONDENCE

Future reference, Badger strain, Paradigm change, Protecting the work 278

NEWS AND VIEWS

Cerebellar modelling (R R Llinás) **Grand bang** (P Davies), **Dominance hierarchies among worker ants** (R M May), **Malaria vaccine protects offspring** (F E G Cox), **Continental swells** (P J Smith), **A revision of the globular cluster metallicity scale** (B F Madore), **Cortical effects of monocular deprivation:** (J A Movshon), **Global warming?** (T M L Wigley, P D Jones, P M Kelly), **Why mammalian gametes don't mix** (J M Bedford), **Understanding natural killer cells** (P Beverley) 279

BOOK REVIEWS 361

Sir Joseph Banks (*by* **C Lyte) and The Journal of Joseph Banks in the Endeavour 1768–1771** AH Dupree, **Some Strangeness in the Proportion: A Centennial Symposium to Celebrate the Achievements of Albert Einstein (H Woolf,** ***ed.*****)** Paul Davies, **The Development of the Vertebrate Limb (*****by*** **JR Hinchliffe and DR Johnson)** Julian Lewis, **Immobilized Enzymes in Analytical and Clinical Chemistry (*****by*** **PW Carr and LD Bowers)** and **Immobilized Enzymes (*****by*** **MD Trevan)** Peter Dunnill, **Biological Monitoring for Environmental Effects (DL Worf,** ***ed.*****)** F Moriarty, **Scanned Image Microscopy (EA Ash,** ***ed.*****)** PW Hawkes, **Predation and Freshwater Communities (*****by*** **TM Zaret)** Geoffrey Fryer, **Basic Principles of Membrane Transport (*****by*** **SG Schultz)** Ora Kedem, **Handbook of Environmental Isotope Geochemistry, Vol 1, A (P Fritz and J Ch Fontes,** ***eds*****)** Gunter Faure

Cover

Cloned zebra fish photographed by Mr Harry Howard (see page 293)

Nature® (ISSN 0028-0836) is published weekly, except the last week in December, by Macmillan Journals Ltd Annual subscription for USA and Canada US $173 Orders (with remittance) and change of address labels to Macmillan Journals Ltd, Brunel Rd, Basingstoke RG21 2XS, UK Second class postage paid at New York, NY 10010 and additional mailing offices US Postmaster send forms 3579 to *Nature*, 15 East 26 Street, New York, NY 10010

ARTICLES

Demise of the Lemaître cosmological models and objects at large redshift D Tytler 289

Production of clones of homozygous diploid zebra fish (*Brachydanio rerio*) G Streisinger, C Walker, N Dower, D Knauber & F Singer 293

DNA-dependent RNA polymerase II of plant origin transcribes viroid RNA into full-length copies H-R Rackwitz, W Rohde & HL Sanger 297

LETTERS TO NATURE

Millimetre observations of quasar Q0420–388 WA Sherwood, GV Schultz & E Kreysa 301

A period–luminosity relation for Mira variables in the Large Magellanic Cloud IS Glass & T Lloyd Evans 303

Modulation of the Earth's electric field by cosmic radiation R Markson 304

Neutron-capture nucleosynthesis of nature's rarest stable isotope H Beer & RA Ward 308

Formation of helium platelets in molybdenum JH Evans, A van Veen & LM Caspers 310

Atmospheric transport of continentally derived lipids to the tropical North Pacific RB Gagosian, ET Peltzer & OC Zafiriou 312

Floristic changes indicate a cooling climate in the Eocene of southern England ME Collinson, K Fowler & MC Boulter 315

A new assessment of *Rhyniella*, the earliest known insect, from the Devonian of Rhynie, Scotland P Whalley & EA Jarzembowski 317

The role of GABAergic inhibition in the cortical effects of monocular deprivation AM Sillito, JA Kemp & C Blakemore 318

Morphine and δ-opiate agonists locally stimulate *in vivo* dopamine release in cat caudate nucleus MF Chesselet, A Chèramy, TD Reisine & J Glowinski 320

Heavy meromyosin cross-links thin filaments in striated muscle myofibrils J Borejdo & A Oplatka 322

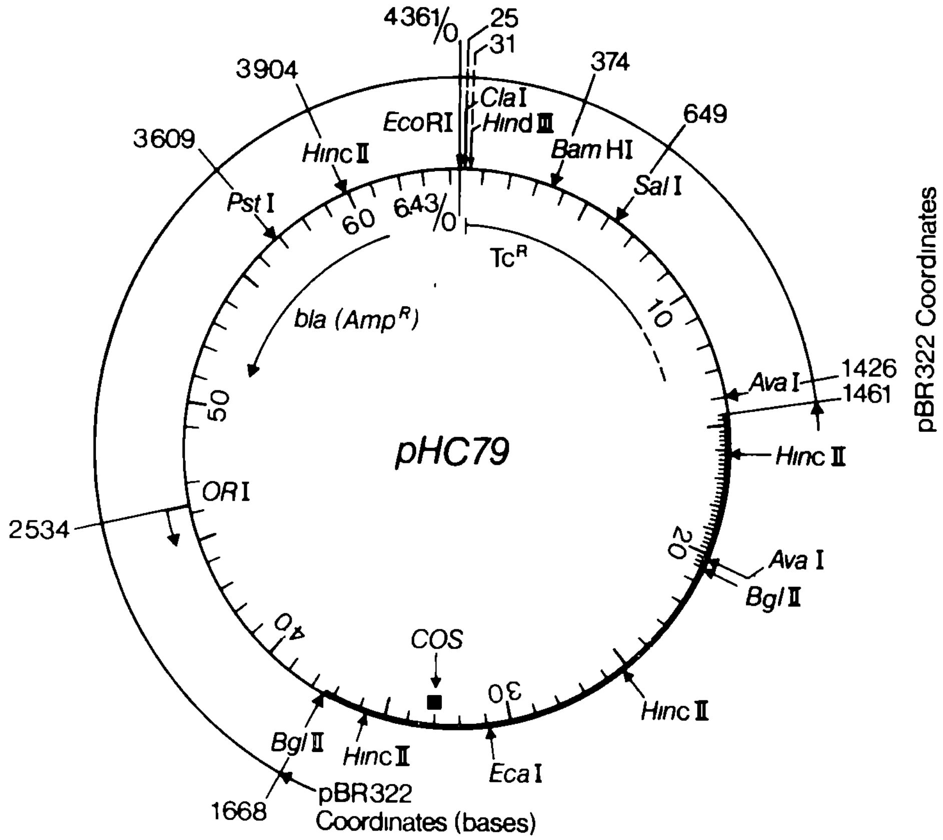
Thromboxane and prostacyclin release from ischaemic myocardium in relation to arrhythmias SJ Coker, JR Parratt, IMcA Ledingham & IJ Zeitlin 323

Sodium current depression by lidocaine and quinidine in isolated ventricular cells KS Lee, JR Hume, W Giles & AM Brown 325

Calcitonin selectively stimulates 25-hydroxyvitamin D_3-1α-hydroxylase in proximal straight tubule of rat kidney H Kawashima, S Torikai & K Kurokowa 327

Contents continued overleaf

Cloning Vectors from Boehringer Mannheim



Hohn, B & Collins, J (1980) Gene 11, 291-298

	Cat No
Deoxyribonucleic acid, Cosmid pHC 79 from Escherichia coli	567 795
Deoxyribonucleic acid, Plasmid pBR 322 (Plasmid pBR 322 DNA) from Escherichia coli	481 238
Deoxyribonucleic acid, Plasmid Col E1 (Plasmid Col E1 - DNA)	208 370
Deoxyribonucleic acid, Plasmid Lambda dv1 (Plasmid Lambda dv1 - DNA)	208 345

and further 150 biochemicals for molecular biological research are from stock available

please contact your local representative or

mannheim boehringer

Boehringer Mannheim GmbH
Biochemica
P O Box 310120
D-6800 Mannheim
W Germany

Circle No 27 on Reader Enquiry Card

MISCELLANY

Erratum 358
Product review xxvii
Classified advertising xxxii

nature

is Available in
MICROFORM

For Complete Information
customers should write:

University Microfilms
International

Dept. F.A.
300 North Zeeb Road
Ann Arbor, MI 48106
U.S.A.

Dept. F.A.
18 Bedford Row
London, WC1R 4EJ
England

Circle No.55 on Reader Enquiry Card.

Annual Subscription including Index (51 issues)

UK & Eire		£75
USA & Canada	Surface only	US$173
Belgium	Airspeed only	BF5695
West German	Airspeed only	DM360
Netherland	Airspeed only	G390
Switzerland	Airspeed only	SF330
Rest of Europe	Airspeed only	£85
Rest of World	Surface	£85
Rest of World*	Airmail	£130

(not USA, Canada & Europe)

Personal subscription rates are available in some countries to subscribers paying by personal cheque or credit card.

Details can be obtained from:
UK: *Nature* Promotions Department,
4 Little Essex Street, London WC2R 3LF
Telephone: Jonathan Earl: 01-836 6633
USA: *Nature*, 15 East 26 Street, New York, NY 10010
Credit card orders only (in USA & Canada):
Call toll-free: (800) 824-7888 (Operator 246)
In California: (800) 857-7777 (Operator 246)

Back issues: (Post-paid) UK, £2.00; USA & Canada, US$6.00 (surface), US$8.50 (air); Rest of World, £2.50 (surface), £3.50 (air).

***Nature* Binders:**
Single Binders: UK, £3.50; Rest of World, $8.50;
Set(s) of Binders: UK, £9.50; Rest of World, $22.50

***Nature* Annual Indexes (1971-1980):**
Prices (post-paid):
UK, £5.00 each; Rest of World, $10.00

***Nature* First Issue Facsimile:**
UK, 75p; Rest of World (surface), $1.50, (air), £2.00

***Nature* Directory of Biologicals**
(Hardcover edition) *(to be published 1981)*
UK, £18.50; Rest of World, $45.00

***Nature* Wallcharts:**
How Britain Runs its Science; How the US Runs its Science; UK, £3.50; USA, $7.95 each

***Nature* in Microform:**
For Information:
UK: UMI, 18 Bedford Row, London WC2R 4EJ
USA: 300 North Zeeb Road, Ann Arbor, Mi 48106

Orders (with remittance) to:
Nature, Macmillan Journals Ltd.,
Brunel Road, Basingstoke, Hants, RG21 2XS,
England

Maintenance of blood pressure by the renin–angiotensin system in normal man GA MacGregor, ND Markandu, JE Roulston, JC Jones & JJ Morton 329

Active immunization and passive transfer of resistance against sporozoite-induced malaria in infant mice AU Orjih, AH Cochrane & RS Nussenzweig 331

Inducer T lymphocytes synthesize a factor that stimulates proliferation of cloned mast cells G Nabel, SJ Galli, AM Dvorak, HF Dvorak & H Cantor 332

***In vivo* effect of anti-asialo GM1 antibody on natural killer activity** M Kasai, T Yoneda, S Habu, Y Maruyama, K Okumura & T Tokunaga 334

Interleukin-2 augments natural killer cell activity CS Henney, K Kuribayashi, DE Kern & S Gillis 335

Synthesis and secretion of mouse immunoglobulin chains from *Xenopus* oocytes G Valle, J Besley & A Colman 338

Transcriptional regulation of mouse liver metallothionein-I gene by glucocorticoids LJ Hager & RD Palmiter 340

Unilateral internuclear gene transfer and cell differentiation in *Schizophyllum* RF Gaber & TJ Leonard 342

A small RNA complementary to an intervening sequence is produced late in SV40 infection R Campos, S Jovanovich & LP Villarreal 344

Identification of the SV40 agnogene product: a DNA binding protein G Jay, S Nomura, CW Anderson & G Khoury 346

Mutagenesis *in vitro* by depurination of ϕ X174 DNA TA Kunkel, CW Shearman & LA Loeb 349

Sym plasmid of *Rhizobium trifolii* expressed in different rhizobial special and *Agrobacterium tumefaciens* PJJ Hooykaas, AAN van Brussel, H. den Dulk-Ras, GMS van Slogteren & RA Schilperoort 351

Recombinant plasmids carrying nitrogen fixation genes from *Rhizobium japonicum* H Hennecke 354

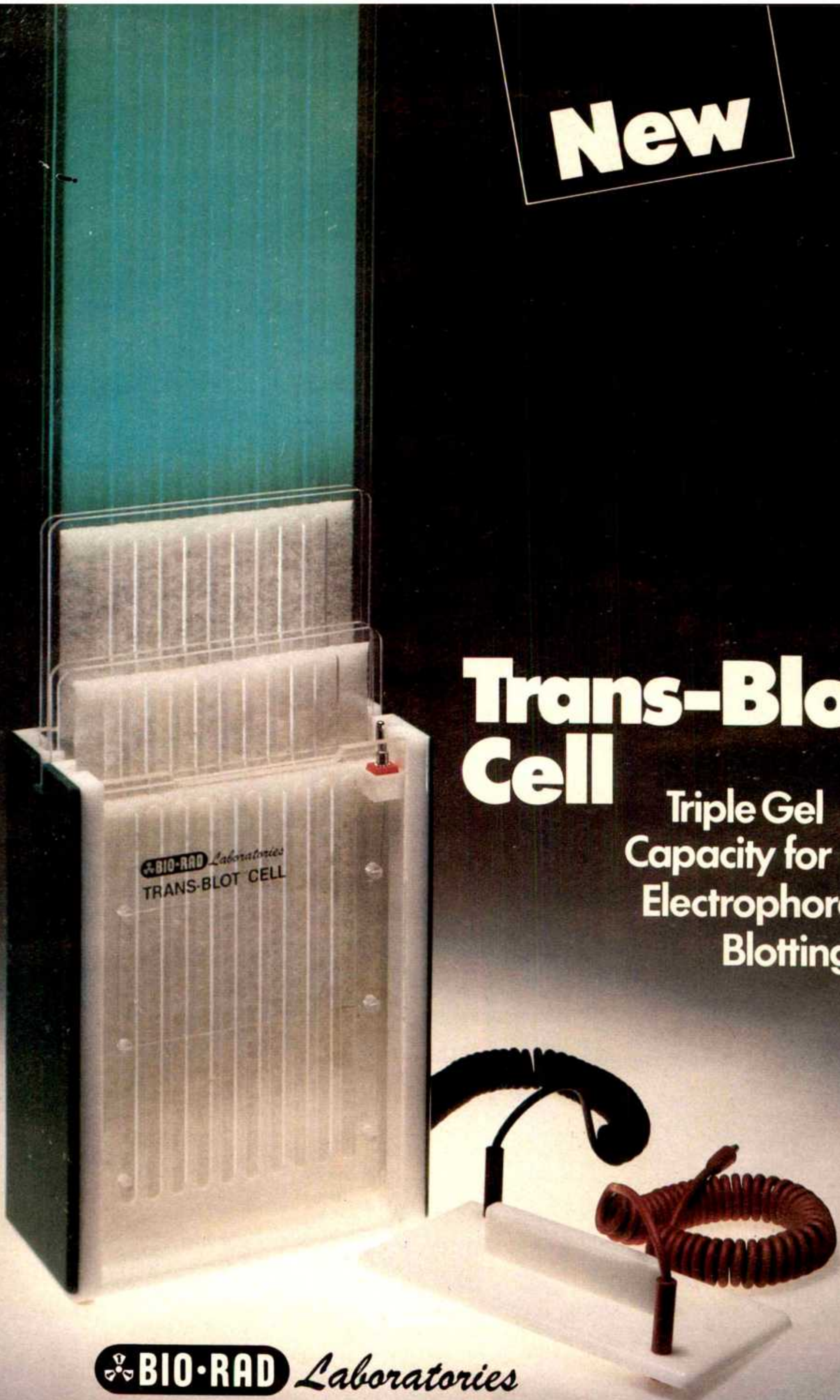
Synthesis and characterization of $(2'\text{-}5')\mathrm{ppp3'dA(p3'dA)}_n$, an analogue of $(2'\text{-}5')\mathrm{pppA(pA)}_n$ P Doetsch, JM Wu, Y. Sawada & RJ Suhadolnik 355

MATTERS ARISING

Polygon patterns on Europa IJ Smalley 359
Reply to Smalley DC Pieri 359

Low-angle X-ray scattering of chromatin JA Subirana 360
Reply to Subirana JP Langmore & C Schutt 360

Stability of loess in light of the inactive particle theory AJ Lutenegger 360
Reply to Lutenegger IJ Smalley 360



New

Trans-Blot™ Cell

Triple Gel Capacity for Electrophoretic Blotting

BIO·RAD *Laboratories*

2200 Wright Avenue • Richmond, CA 94804 USA • Phone (415) 234-4130
Also in Rockville Centre, N.Y., Australia, Austria, England, Germany,
Italy, Japan, The Netherlands, and Switzerland.

See us at the Biochemistry Meeting
Booths N-13, 14, 15

nature

Volume 291 No. 5814 4 June 1981

[illegible] Library

Madness, responsibility and the law 367
Nuclear politics 368

NEWS

NIH censure for Dr Martin Cline
Social scientists take over ANZAAS forum
Mitterand's new ministers
French face energy questions
Corporate largesse
Yugoslovian university unrest
US security worries
UK agricultural research
French to make Arabsat
Origin of Species
Huxley on *Nature* 369

CORRESPONDENCE

Apes or angels?/It wasn't me/A relative genius 374

NEWS AND VIEWS

Geophysical controversy in the Gulf of Aden (R B Whitmarsh) **The transmission and control of gonorrhea** (R M May) **150 years of the North Magnetic Pole** (D R Barraclough, S R C Malin) **The treatment of inborn errors of the urea cycle** (I Smith) **Late-glacial climatic changes** (P Moore) 375

BOOK REVIEWS

Wholeness and the Implicate Order (*by* D Bohm) A Shimony; **Muscle Contraction: Its Regulatory Mechanisms (S Ebashi, *ed.*)** C R Bagshaw; **Experimental Methods in Polymer Chemistry (*by* J F Rabek)** H Mark; **Evolution of Physical Oceanography (B A Warren and C Wunsch, *eds*)** K F Bowden; **Statistical Mechanics of the Liquid Surface (*by* C A Croxton)** P Schofield; **The Iris in Eighteenth-century Physiology (*by* R G Mazzolini)** J C Dean Hart 435

MISCELLANY

100 years ago 378
Books received 439
Announcements 440
New on the market xxi
Classified advertising xxiii

Cover

The distribution of opiate receptors, labelled with tritiated naloxone, in the caudate nucleus of the brain. A computer was used to analyse grain density of LKB tritium-sensitive film (see page 415)

Nature® (ISSN 0028-0836) is published weekly, except the last week in December, by Macmillan Journals Ltd. Annual subscription for USA and Canada US $173. Orders (with remittance) and change of address labels to: Macmillan Journals Ltd, Brunel Rd, Basingstoke RG21 2XS, UK. Second class postage paid at New York, NY 10010 and additional mailing offices. US Postmaster send form 3579 to: *Nature*, 15 East 26 Street, New York, NY 10010.

The evolution of metabolic cycles J E Baldwin & H Krebs 381

ARTICLES

Microwave OH master emission in the circumstellar envelopes of late-type stars P F Bowers, K J Johnston & J H Spencer 382
Molecular analysis of the *unc*–54 myosin heavy-chain gene of *Caenorhabditis elegans* A R MacLeod, J Karn & S Brenner 386

LETTERS TO NATURE

Why is the Universe open? D Lindley 391
IR variability of SS433 R M Catchpole, I S Glass, B S Carter & G Roberts 392
Thermal continuum emission from magnetic rotators in quasars W H Sorrell 394
An Antarctic iron meteorite contains preterrestrial impact-produced diamond and lonsdaleite R S Clarke Jr, D E Appleman & D R Ross 396
Excitation of the green line in the night airglow R D Kenner, E A Ogryzlo & P T Wassell 398
Porous titania glass as a photocatalyst for hydrogen production from water L Kruczynski, H D Gesser, C W Turner & E A Speers 399
Scanning electron-stimulated desorption microscopy of model biological surfaces H F Dylla, J H Abrams, C T Hovland & J G King 401
Comparisons of sulphide deposits from the East Pacific Rise and Cyprus E Oudin, P Picot & G Pouit 404
Are high rates of sulphate reduction associated with anaerobic oxidation of methane? A H Devol & S I Ahmed 407
Albumin systematics of the extinct mammoth and Tasmanian wolf J M Lowenstein, V M Sarich & B J Richardson 409
Intense exercise, bone structure and blood calcium levels in vertebrates J A Ruben & A F Bennett 411
Somatic and behavioural postnatal effects of fetal injections of nerve growth factor antibodies in the rat L Aloe, C Cozzari, P Calissano & R Levi-Montalcini 413
Mosaic distribution of opiate receptors, parafascicular projections and acetylcholinesterase in rat striatum M Herkenham & C B Pert 415
Cell-surface antigen distinguishes sensory and autonomic peripheral neurones from central neurones T Vulliamy, S Rattray & R Mirsky 418

0028-0836/81/23000C-02$01.00

Contents continued overleaf

SAFE

in the knowledge of Barr & Stroud technology at work

A commonplace scene on a busy road, but safer if Barr & Stroud fibre optic traffic signs were installed.

Our signs allow virtually no misleading phantom images even in the most difficult conditions.

We are a leading light in the fibre optics field, where our experience spans a wide spectrum of equipment, systems and applications – in computers, process control, medicine, colour detection, automatic inspection, as well as traffic signs.

Barr & Stroud's leadership and innovative skills in today's technologies are at the service of science and industry.

We have proved it in fibre optics; we have proved it in other areas. Because of our serious commitment to research and development, engineers and designers everywhere can rely on our expertise to solve their most complex problems.

Barr & Stroud
make fibre optics work

Barr & Stroud Limited,
Caxton Street, Anniesland, Glasgow G13 1HZ.
Melrose House,
4-6 Savile Row, London W1X 1AF.

PILKINGTON GROUP

Circle No. 05 on Reader Enquiry Card.

• How Britain runs its Science

• How the US runs its Science

These two new *Nature* wallcharts are now available.

The 2 x 3ft guides offer a comprehensive survey of the administration of science in Britain and the US, showing the interlocking relationship between science, engineering and medical establishments

Designed as fold-out sheets, the wallcharts can be carried easily to meetings or hung as posters for permanent reference

The new 1981 wallcharts are available at £3 50 or $7 95 each from *Nature's* offices in London, Basingstoke and New York

Annual Subscription including Index (51 issues)

UK & Eire		£75
USA & Canada	Surface only	US$173
Belgium	Airspeed only	BF5695
West German	Airspeed only	DM360
Netherland	Airspeed only	G390
Switzerland	Airspeed only	SF330
Rest of Europe	Airspeed only	£85
Rest of World	Surface	£85
Rest of World*	Airmail	£130

**(not USA, Canada & Europe)*

Personal subscription rates are available in some countries to subscribers paying by personal cheque or credit card

Details can be obtained from
UK *Nature* Promotion Department,
4 Little Essex Street, London WC2R 3LF
Telephone Jonathan Earl 01-836 6633

USA *Nature*, 15 East 26 Street, New York, NY 10010
Credit card orders only (in USA & Canada)
Call toll-free (800) 824-7888 (Operator 246)
In California (800) 857-7777 (Operator 246)

Back issues (Post-paid) UK, £2 00, USA & Canada, US$6 00 (surface), US$8 50 (air), Rest of World, £2 50 (surface), £3 50 (air)

***Nature* Binders**
Single Binders UK, £3 50, Rest of World, $8 50,
Set(s) of Binders UK, £9 50, Rest of World, $22 50

***Nature* Annual Indexes (1971-1980)**
Prices (post-paid)
UK, £5 00 each, Rest of World, $10 00

***Nature* First Issue Facsimile**
UK, 75p, Rest of World (surface), $1 50, (air), £2 00

***Nature* Directory of Biologicals**
(Hardcover edition) *(to be published 1981)*
UK, £18 50, Rest of World, $45 00

***Nature* Wallcharts·**
How Britain Runs its Science, How the US Runs its Science, UK, £3 50, USA, $7 95 each

***Nature* in Microform:**
For Information
UK UMI, 18 Bedford Row, London WC2R 4EJ
USA 300 North Zeeb Road, Ann Arbor, Mi 48106

Orders (with remittance) to
Nature, Macmillan Journals Ltd ,
Brunel Road, Basingstoke, Hants, RG21 2XS,
England

A neuronal cell-surface antigen is found in the CNS but not in peripheral neurones J Cohen & S Y Selvendran 421

Non-random openings and concentration-dependent lifetimes of glutamate-gated channels in muscle membrane K A F Gration, J J Lambert, R R Ramsey & P N R Usherwood 423

Sodium channels need not open before they inactivate R Horn, J Patlak & C F Stevens 426

K^+ channels close more slowly in the presence of external K^+ and Rb^+ R P Swenson Jr & C M Armstrong 427

Transcription of single-copy vitellogenin gene of *Xenopus* involves expression of middle repetitive DNA G U Ryffel, D B Muellener, T Wyler, W Wahli & R Weber 429

Isolation of virus from brain after immunosuppression of mice with latent herpes simplex L Kastrukoff, C Long, P C Doherty, Z Wroblewska & H Koprowski 432

Disorientation of inexperienced young pigeons after transportation in total darkness W Wiltschko & R Wilstschko 433

GUIDE TO AUTHORS

Authors should be aware of the diversity of *Nature*'s readership and should strive to be as widely understood as possible

Review articles should be accessible to the whole readership Most are commissioned, but unsolicited reviews are welcome (in which case prior consultation with the office is desirable)

Scientific articles are research reports whose conclusions are of general interest or which represent substantial advances of understanding The text should not exceed 3,000 words and six displayed items (figures plus tables) A draft abstract (50 words) is required

Letters to *Nature* are ordinarily 1,000 words long with no more than four displayed items The first paragraph (not exceeding 150 words) should say what the letter is about, why the study it reports was undertaken and what the conclusions are

Matters arising are brief comments (up to 500 words) on articles and letters recently published in *Nature*

Manuscripts may be submitted either to London or New York (decisions being made only in London) Manuscripts should be typed (double spacing) on one side of the paper only Three copies are required, each accompanied by copies of lettered artwork No title should exceed 80 characters in length Reference lists, figure legends, etc should be on separate sheets, all of which should be numbered Abbreviations, symbols, units, etc should be identified on one copy of the manuscript at their first appearance

References should appear sequentially indicated by superscripts in the text and should be abbreviated according to the *World List of Scientific Periodicals*, fourth edition (Butterworth 1963–65) The first and last page numbers of each reference should be cited References to books should clearly indicate the publisher and the date and place of publication Unpublished articles should not be formally referred to unless accepted or submitted for publication, but may be mentioned in the text

Each piece of artwork should be clearly marked with the author's name and the figure number Original artwork should be unlettered Suggestions for cover illustrations are welcome Original artwork (and one copy of the manuscript) will be returned when manuscripts cannot be published

[illegible]

An explanation why Swissair has no need to introduce a King Lounge Class, a Flopper-Hopper Class, or Golden Edelweiss Service.

nature

Volume 291 No. 5815 11 June 1981

What future for postdoctoral fellows? 441
Change wanted 442
Some silver lining 442

NEWS

Anxiety about postdoctoral prospects 443
Low pay, indignities and uncertainty
US minerals
National Cancer Institute
Vive la différence!
US astronomy
Uranium enrichment
UK nuclear power
French hiatus
Yugoslavian nuclear power
Research in Poland
London University
Psychiatrist on trial

CORRESPONDENCE

Post Office inertia/Bovine units/Look harder/The fact is 448

NEWS AND VIEWS

Overfishing: is there a solution? (A J Southward)
Collagen evolution (E Solomon, & K S E Cheah)
Polymers by synchrotron radiation (P Calvert)
Atom-specific bond breakage (M L Knotek)
Neuropeptides· do they integrate body and brain? (S D Iversen) 449

BOOK REVIEWS

A Balance Sheet on Burt (H Beloff, *ed*) J M Thoday, **Radio Physics of the Sun (M R Kundu & T E Gergely, *eds*)** D Layzer, **The Search for Gravity Waves (*by* P C W Davies)** W D Allen, **Hormonal Control of Lactation (*by* A T Cowie *et al*)** A S McNeilly, **Monoclonal Antibodies. Hybridomas: A New Dimension in Biological Analyses (R H Kennett *et al*, *eds*)** C M Croce, **British Anthozoa (*by* R L Manuel)** C M Yonge 517

MISCELLANY

100 years ago 454
Books received 522
Product review: Molecular biology xxiii
Classified advertising xxvii

Cover

An electron micrograph showing molecular detail of the quaterrylene $C_{40}H_{20}$, taken with the Cambridge University 600-kV high resolution electron microscope (see page **481**)

Nature® (ISSN 0028 0836) is published weekly, except the last week in December, by Macmillan Journals Ltd Annual subscription for USA and Canada US $173 Orders (with remittance) and change of address labels to Macmillan Journals Ltd, Brunel Rd, Basingstoke RG21 2XS, UK Second class postage paid at New York, NY 10010 and additional mailing offices US Postmaster send form 3579 to *Nature*, 15 East 26 Street, New York, NY10010

The traditional and a new version of the mouse *H-2* complex (a review) J Klein, A Juretić, C N Baxevanis & Z A Nagy 455

ARTICLES

Lead isotopic composition of Hercynian granitic K-feldspars constrains continental genesis A M Vitrac, F Albarede & C J Allègre 460
Only one of two closely related yeast suppressor tRNA genes contains an intervening sequence M V Olson, G S Page, A Sentenac, P W Piper, M Worthington, R B Weiss & B D Hall 464

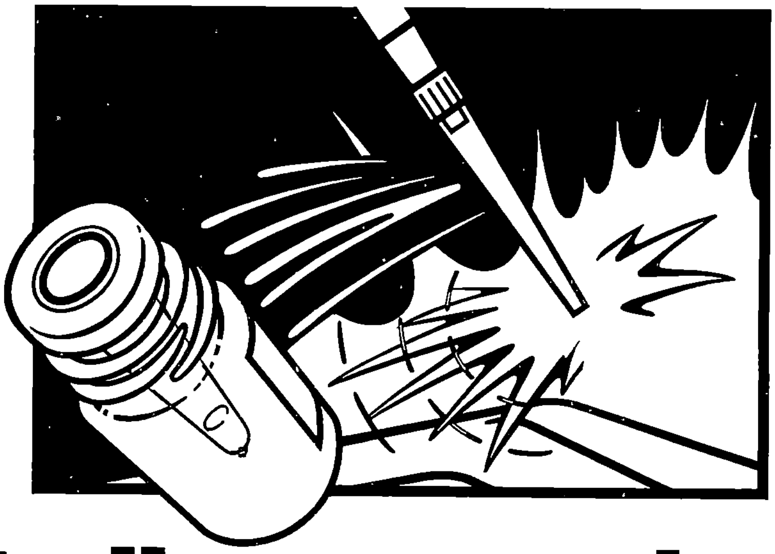
LETTERS TO NATURE

Baryosynthesis and the origin of galaxies J D Barrow & M S Turner 469
Impact fusion and the field emission projectile E R Harrison 472
British shore platforms and ice-sheets J B Sissons 473
Will the Sahelian drought end in 1985? H Faure & J-Y Gac 475
Geological evidence against major displacement in the Nares Strait R L Christie, P R Dawes, T Frisch, A K Higgins, J M Hurst, J W Kerr & J S Peel 478
Molecular detail in electron micrographs of quaterrylene $C_{40}H_{20}$ D J Smith & J R Fryer 481
Model protocells photochemically reduce carbonate to organic carbon C Folsome & A Brittain 482
Structurally preserved angiosperm flowers from the Upper Cretaceous of southern Sweden E M Friis & A Skarby 484
Congruent shifts in sand eel abundance in western and eastern North Atlantic ecosystems K Sherman, C Jones, L Sullivan, W Smith, P Berrien & L Ejsymont 486
Monocyclic β-lactam antibiotics produced by bacteria R B Sykes, C M Cimarusti, D P Bonner, K Bush, D M Floyd, N H Georgopapadakou, W H Koster, W C Liu, W L Parker, P A Principe, M L Rathnum, W A Slusarchyk, W H Trejo & J S Wells 489
Vasopressor receptor antagonist prevents behavioural effects of vasopressin M Le Moal, G F Koob, L Y Koda, F E Bloom, M Manning, W H Sawyer & J Rivier 491
Chronic morphine administration increases the apparent number of α_2-adrenergic receptors in rat brain M Hamburg & J F Tallman 493
Contralateral denervation causes enhanced transmitter release from frog motor nerve terminals A A Herrera & A D Grinnell 495

0028-0836/81/24000C-02$01 00

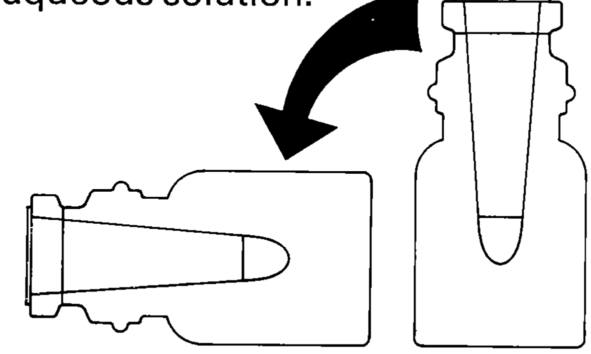
Contents continued overleaf

Accidents will happen...



...but all may not be lost!

Amersham's aqueous solutions* of ^{32}P-nucleotides at 10mCi/ml are supplied in "Duoseal V-vials". In an experiment designed to simulate a laboratory accident, 10 open "Duoseal V-vials" were knocked over, each containing 100μl of aqueous solution.



In all cases the surface tension prevented any loss from the specially designed V-insert.

Other advantages of the V-insert

- will accept standard pipette tips
- minimum loss of product, <2% per week lost by absorption
- is not 'wetted' by the solution
- will fit caps from small centrifuge tubes

Aqueous solutions of ^{32}P-nucleotides available weekly from Amersham in "Duoseal V-vials".

Call Amersham or your local representative for further details.

* contains 5mM 2-mercaptoethanol

Amersham International Limited, Amersham, England Telephone 024-04-4444
In the U S A and Canada, Amersham Corporation, Illinois 60005 Telephone 312-364 7100 and 800 323-9750 (Tollfree)
In W Germany, Amersham Buchler GmbH & Co KG, Braunschweig Telephone 05307-4691

Circle No 28 on Reader Enquiry Card

Amersham

• **How Britain runs its Science**

• **How the US runs its Science**

These two new *Nature* wallcharts are now available.

The 2 x 3ft guides offer a comprehensive survey of the administration of science in Britain and the US, showing the interlocking relationship between science, engineering and medical establishments

Designed as fold-out sheets, the wallcharts can be carried easily to meetings or hung as posters for permanent reference

The new 1981 wallcharts are available at £3 50 or $7 95 each from *Nature's* offices in London, Basingstoke and New York

Annual Subscription including Index (51 issues)

UK & Eire		£75
USA & Canada	Surface only	US$173
Belgium	Airspeed only	BF5695
West German	Airspeed only	DM360
Netherland	Airspeed only	G390
Switzerland	Airspeed only	SF330
Rest of Europe	Airspeed only	£85
Rest of World	Surface	£85
Rest of World*	Airmail	£130

**(not USA, Canada & Europe)*

Personal subscription rates are available in some countries to subscribers paying by personal cheque or credit card

Details can be obtained from
UK *Nature* Promotion Department,
4 Little Essex Street, London WC2R 3LF
Telephone Jonathan Earl 01-836 6633

USA *Nature*, 15 East 26 Street, New York, NY 10010
Credit card orders only (in USA & Canada)
Call toll-free (800) 824-7888 (Operator 246)
In California (800) 857-7777 (Operator 246)

Back issues (Post-paid) UK, £2 00, USA & Canada, US$6 00 (surface), US$8 50 (air), Rest of World, £2 50 (surface), £3 50 (air)

***Nature* Binders**
Single Binders UK, £3 50, Rest of World, $8 50,
Set(s) of Binders UK, £9 50, Rest of World, $22 50

***Nature* Annual Indexes (1971-1980)**
Prices (post-paid)
UK, £5 00 each, Rest of World, $10 00

***Nature* First Issue Facsimile**
UK, 75p, Rest of World (surface), $1 50, (air), £2 00

***Nature* Directory of Biologicals**
(Hardcover edition) *(to be published 1981)*
UK, £18 50, Rest of World, $45 00

***Nature* Wallcharts**
How Britain Runs its Science, How the US Runs its Science, UK, £3 50, USA, $7 95 each

***Nature* in Microform**
For Information
UK UMI, 18 Bedford Row, London WC2R 4EJ
USA 300 North Zeeb Road, Ann Arbor, Mi 48106

Orders (with remittance) to
Nature, Macmillan Journals Ltd ,
Brunel Road, Basingstoke, Hants, RG21 2XS, England

Ca-dependent K channels with large unitary conductance in chromaffin cell membranes A Marty 497

Monoclonal antibodies against human fibroblast interferon H K Hochkeppel, U Menge & J Collins 500

Production of both macrophage activating and inhibiting activities by a murine T-lymphocyte hybridoma C M Jones, J A Braatz & R B Herberman 502

Synthesis of hepatitis B surface and core antigens in *E. coli* J C Edman, R A Hallewell, P Valenzuela, H M Goodman & W J Rutter 503

Lack of actin III in fibrillar flight muscle of flightless *Drosophila* mutant *raised* A B Lang, C Wyss & H M Eppenberger 506

An adenovirus mutant defective in splicing RNA from early region 1A D Solnick 508

***In vitro* synthesis of polypeptides encoded by the long terminal repeat region of mouse mammary tumour virus DNA** C Dickson, R Smith & G Peters 511

Species variation in the specificity of ribulose bisphosphate carboxylase/oxygenase D B Jordan & W L Ogren 513

Modification of calmodulin function by enzymatic carboxyl methylation C Gagnon, S Kelly, V Manganiello, M Vaughan, C Odya, W Strittmatter, A Hoffman & F Hirata 515

GUIDE TO AUTHORS

Authors should be aware of the diversity of *Nature*'s readership and should strive to be as widely understood as possible

Review articles should be accessible to the whole readership Most are commissioned, but unsolicited reviews are welcome (in which case prior consultation with the office is desirable)

Scientific articles are research reports whose conclusions are of general interest or which represent substantial advances of understanding The text should not exceed 3,000 words and six displayed items (figures plus tables) A draft abstract (50 words) is required

Letters to *Nature* are ordinarily 1,000 words long with no more than four displayed items The first paragraph (not exceeding 150 words) should say what the letter is about, why the study it reports was undertaken and what the conclusions are

Matters arising are brief comments (up to 500 words) on articles and letters recently published in *Nature*

Manuscripts may be submitted either to London or New York (decisions being made only in London) Manuscripts should be typed (double spacing) on one side of the paper only Three copies are required, each accompanied by copies of lettered artwork No title should exceed 80 characters in length Reference lists, figure legends, etc should be on separate sheets, all of which should be numbered Abbreviations, symbols, units, etc should be identified on one copy of the manuscript at their first appearance

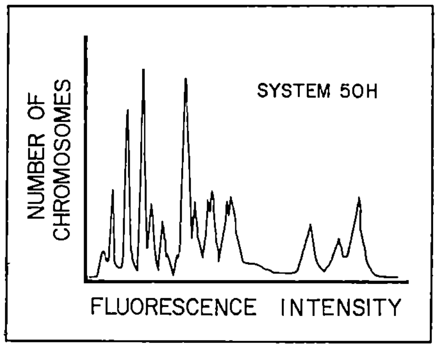
References should appear sequentially indicated by superscripts in the text and should be abbreviated according to the *World List of Scientific Periodicals*, fourth edition (Butterworth 1963–65) The first and last page numbers of each reference should be cited References to books should clearly indicate the publisher and the date and place of publication Unpublished articles should not be formally referred to unless accepted or submitted for publication, but may be mentioned in the text

Each piece of artwork should be clearly marked with the author's name and the figure number Original artwork should be unlettered Suggestions for cover illustrations are welcome Original artwork (and one copy of the manuscript) will be returned when manuscripts cannot be published

No other commercially available cell sorter and analyzer approaches the power, precision and versatility of the Ortho Systems 30/50 The instrument is offered in two basic configurations System 30—a high resolution, 21 parameter analytical flow cytometer, and System 50—which adds cell sorting capabilities to the System 30 flow cytometer Both feature dual lasers as standard equipment for increased analytical power

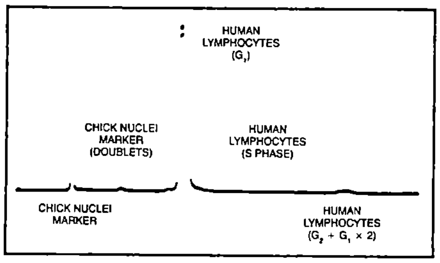
System Components

To appreciate the far-reaching capabilities of these Ortho instruments, consider that the System 50 combines two lasers—a 8 milliwatt helium-neon for ultra high precision scatter measurements and a 5 watt argon type for fluorescence measurements (optional argon-ion, and krypton-ion lasers with U V capability are available), four detectors—two of the photo-multiplier type for right angle scatter, and fluorescent emission detection (S-20 response, 185 nm–850 nm), and two solid state sensors for axial light loss and narrow angle forward scatter a sophisticated optical system incorporating fiber optics a multiplexed multichannel analyzer, with C R T display the signal processor unit and an advanced cell sorter module Additionally, the Model 2150 computerized 8 parameter data handler/sorter controller is available as an option* The 2150 System features simultaneous data acquisition, and real time processing while sorting is in progress



2 Histogram of CHO chromosomes stained for D N A. Ethidium Bromide/Chromomycin A_3

Cytofluorograph System 50H Laser-250mw @ 457nm Red Fluorescence



2 Histogram of human lymphocytes utilizing a chick nuclei marker for standardization and stained for D N A with Propidium Iodide C V = 1 65% (Human Lymphocytes)

System Capabilities

Together, these modules allow the System 50 to provide and analyze morphological information—a significant first for a flow cytometric instrument With four detectors and the ability to display pulse height, pulse area or pulse width analysis for each—a total of 21 distinct measurement parameters can be studied

The System 50 also features two bidimensional regions of interest, dual histogram multi-channel analyzer with cytogram mode, super-sil quartz optics with which the cells are analyzed and easy sample entry

In order to provide customers with complete support, Ortho maintains an Applications Laboratory that is available for confirmation and consultation services Additionally, an international network of service technicians is ready to aid customers should calibration or repairs ever be required

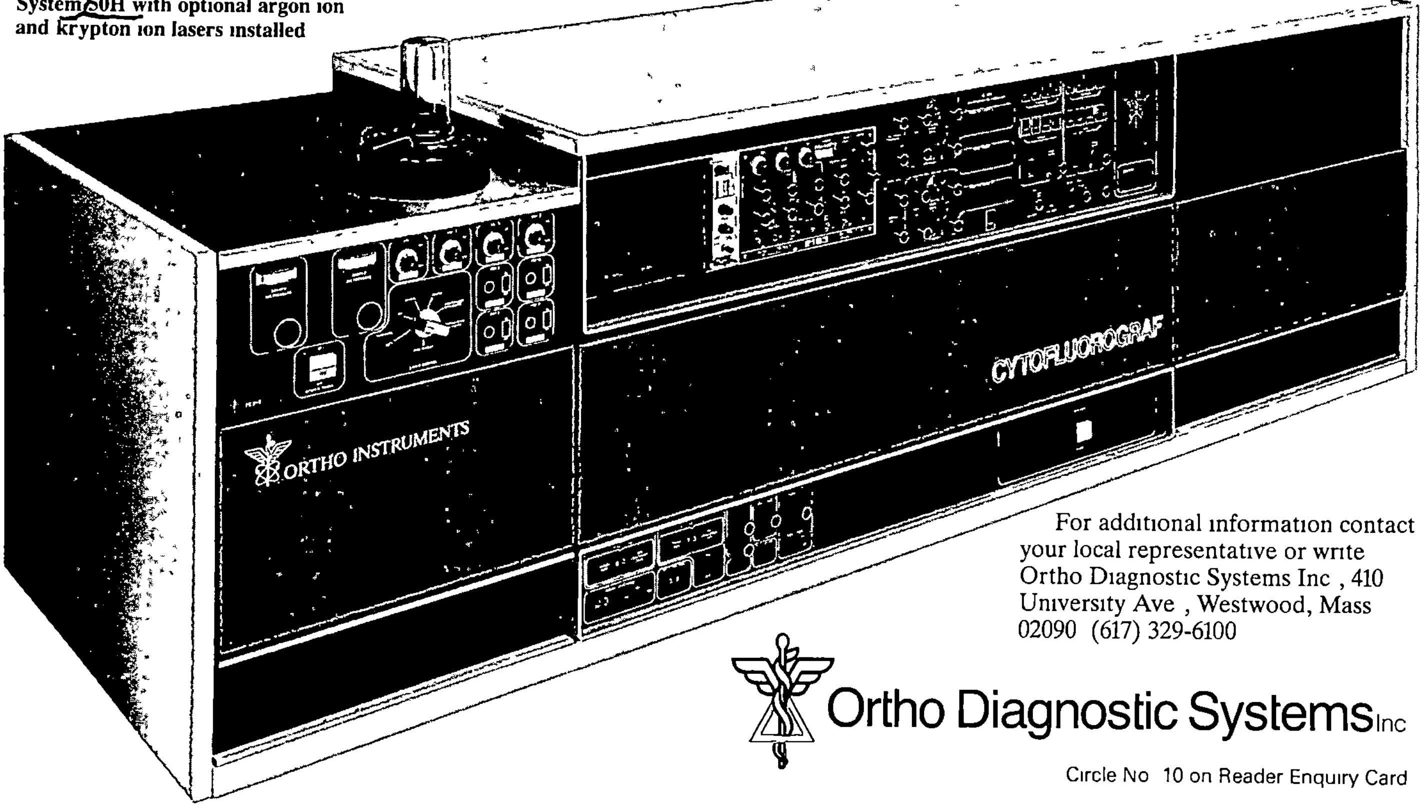
Remarkable Results

The histograms and cytograms displayed were produced by the System 50 Evaluate them And consider the difficulty of deriving this data by any other means The capabilities of this unique Ortho instrument will speak for themselves

State-of-the-art cell sorting and analysis from Ortho

The Cytofluorograf® Systems 30/50

System 50H with optional argon ion and krypton ion lasers installed



For additional information contact your local representative or write Ortho Diagnostic Systems Inc , 410 University Ave , Westwood, Mass 02090 (617) 329-6100

Ortho Diagnostic Systems Inc

Circle No 10 on Reader Enquiry Card

nature

Volume 291 No. 5816 18 June 1981

Making Israel atone for Tamuz 523
Cheque-book academics in trouble 524

NEWS

Hoechst makes deal with Mass. General
NIH plan new overhead calculation
Exploiting research
Yugoslavs to resign
Tropical disease finance
Armadillos fight leprosy
US solar energy
European space agency
All aboard for Halley
European innovation
Soviet scientists
Optical astronomy
Trademarks and patents
Cook's aid surfaces 525

CORRESPONDENCE

Role of US Academy/Questionable results/Film-loop evolves /Had enough too? 530

NEWS AND VIEWS

Gene transfer in plants (R Davies)
Cell sorting, cell differentiation and cyclic AMP (A Tsang)
Scratches on the Palaeolithic record (C S Gamble)
The discovery of free quarks? (L Lyons)
Raman: an eye for combustion studies (D A Greenhalgh)
Amazonian question (D H Tarling)
Opportunism of plant seeds (P D Moore) 531

BOOK REVIEWS

The Expanding Circle: Ethics and Sociobiology (*by* **P Singer**) V Reynolds, **Mental Imagery and Human Memory** (*by* **J T E Richardson**) M Moscovitch, **Chemistry and Biogeochemistry of Estuaries** (**E Olausson and I Cato,** *eds*)J D Burton, **Missing Links** (*by* **J Reader**) and **Lucy: The Beginnings of Humankind** (*by* **D Johanson**) A Hill, **Advances in Infrared Group Frequencies** (*by* **L J Bellamy**) R C Lord, **Chloroplasts** (**J Reinert,** *ed.*) F R Whatley 597

Cover
A temple fresco from Memphis, ancient capital of Egypt, showing a priest afflicted with "horse foot", a typical effect of poliomyelitis The full gene sequence of poliovirus is published this week (see page 547)

Nature® (ISSN 0028-0836) is published weekly, except the last week in December, by Macmillan Journals Ltd Annual subscription for USA and Canada US $173 Orders (with remittance) and change of address labels to Macmillan Journals Ltd, Brunel Rd, Basingstoke RG21 2XS, UK Second class postage paid at New York, NY 10010 and additional mailing offices US Postmaster send form 3579 to *Nature,* 15 East 26 Street, New York, NY10010

Mechanisms of slow postsynaptic potentials (a review) — H C Hartzell 539

ARTICLES

Graphite–magnetite aggregates in ordinary chondritic meteorites — E R D Scott, G J Taylor, A E Rubin, A Okada & K Keil 544
Primary structure, gene organization and polypeptide expression of poliovirus RNA — N Kitamura, B L Semler, P G Rothberg, G R Larsen, C J Adler, A J Dorner, E A Emini, R Hanecak, J J Lee, S van der Werf, C W Anderson & E Wimmer 547
Effects of sleep and arousal on the processing of visual information in the cat — M S Livingstone & D H Hubel 554

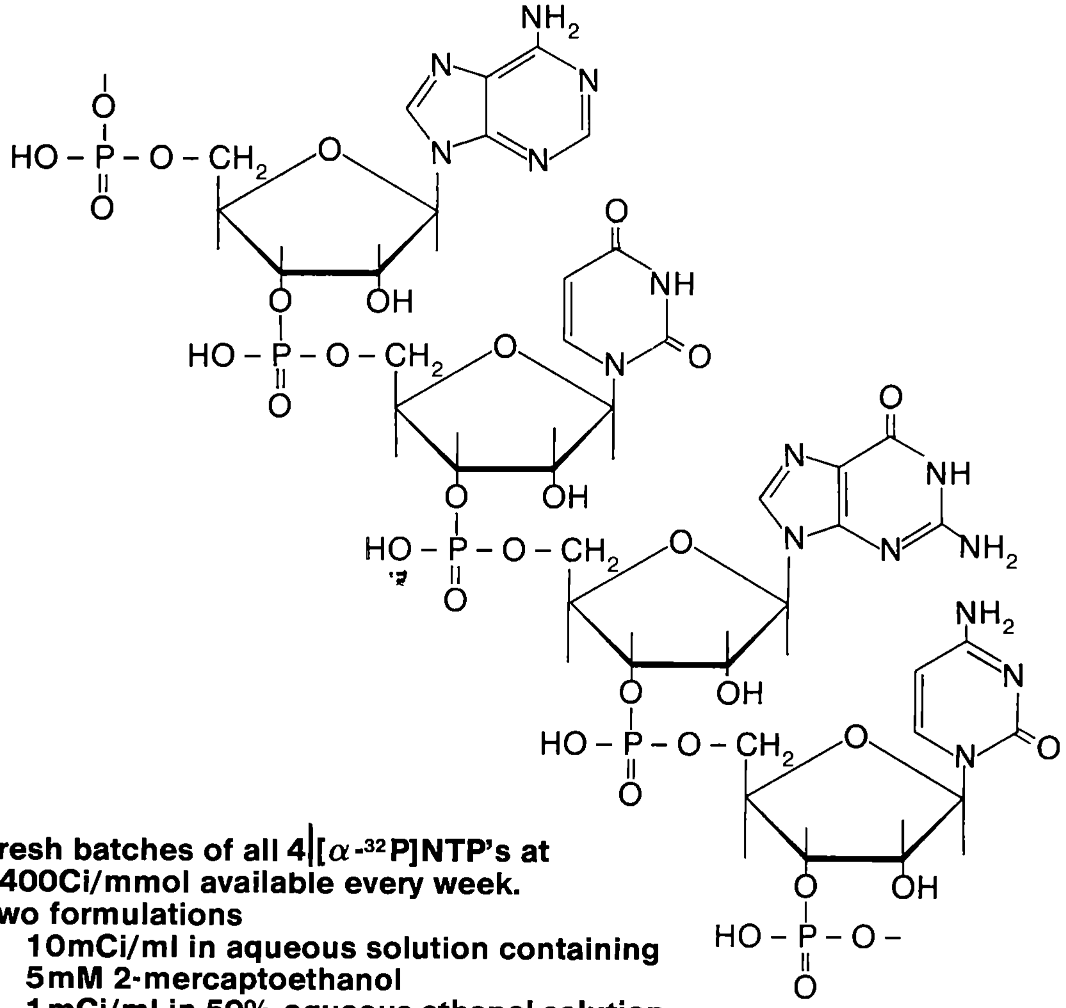
LETTERS TO NATURE

Density power spectrum in the local interstellar medium — J W Armstrong, J M Cordes & B J Rickett 561
Atmospheric nitrous oxide produced by solar protons and relativistic electrons — S S Prasad & E C Zipf 564
Field and structure behaviour of electrolytes — J W Cobble, R C Murray Jr & U Sen 566
Blast dynamics at Mount St Helens on 18 May 1980 — S W Kieffer 568
A model for the origin of the alkaline complexes of Egypt — P de Gruyter & T A Vogel 571
Archaeological evidence for meat-eating by Plio-Pleistocene hominids from Koobi Fora and Olduvai Gorge — H T Bunn 574
Cutmarks made by stone tools on bones from Olduvai Gorge, Tanzania — R Potts & P Shipman 577
Chondroitin sulphate from fossilized antlers — J E Scott & E W Hughes 580
Monoclonal antibodies that distinguish between New World species of *Leishmania* — D McMahon Pratt & J R David 581
Induction of a light requirement during seed development and its ecological consequences — E G Cresswell & J P Grime 583
Depression of net photosynthesis in *Vicia faba* L. exposed to sulphur dioxide and ozone — D P Ormrod, V J Black & M H Unsworth 585
Gene transfer in *Nicotiana rustica* using irradiated pollen — J L Jinks, P D S Caligari & N R Ingram 586
Visual cortex of an anthropoid ape — F Vital-Durand & C Blakemore 588

0028-0836/81/25000C-02$01 00

Contents continued overleaf

Ribonucleoside 5′-[α-^{32}P] triphosphates Straight from the bottle every week! From 7th May 1981



Fresh batches of all 4 [α-^{32}P]NTP's at >400Ci/mmol available every week.
Two formulations

- **10mCi/ml in aqueous solution containing 5mM 2-mercaptoethanol**
- **1mCi/ml in 50% aqueous ethanol solution**

Radiochemical Purity

- All batches analysed by HPLC

Aqueous solutions at 10mCi/ml

- Supplied in "Duoseal V-vials"
- Reduced need for evaporation of the solvent
- Avoids the risk of decomposition during evaporation
- Reduced handling minimizes risk of wastage, spillage and contamination

Please ensure you have your copy of our despatch schedule
Call Amersham or your local representative for details.

Amersham International Limited Amersham England Telephone 024 04 4444
In the USA and Canada Amersham Corporation, Illinois 60005 Telephone 312 364 7100 and 800/323 9750 (Tollfree)
In W Germany Amersham Buchler GmbH & Co KG, Braunschweig. Telephone 05307 4691

Amersham

Circle No 09 on Reader Enquiry Card

- **How Britain runs its Science**
- **How the US runs its Science**

These two new *Nature* wallcharts are now available.

The 2 x 3ft guides offer a comprehensive survey of the administration of science in Britain and the US, showing the interlocking relationship between science, engineering and medical establishments

Designed as fold-out sheets, the wallcharts can be carried easily to meetings or hung as posters for permanent reference

The new 1981 wallcharts are available at £3 50 or $7 95 each from *Nature's* offices in London, Basingstoke and New York

Annual Subscription including Index (51 issues)

UK & Eire		£75
USA & Canada	Surface only	US$173
Belgium	Airspeed only	BF5695
West German	Airspeed only	DM360
Netherland	Airspeed only	G390
Switzerland	Airspeed only	SF330
Rest of Europe	Airspeed only	£85
Rest of World	Surface	£85
Rest of World*	Airmail	£130

**(not USA, Canada & Europe)*

Personal subscription rates are available in some countries to subscribers paying by personal cheque or credit card

Details can be obtained from
UK *Nature* Promotion Department,
4 Little Essex Street, London WC2R 3LF
Telephone Jonathan Earl 01-836 6633

USA *Nature*, 15 East 26 Street, New York, NY 10010
Credit card orders only (in USA & Canada)
Call toll-free (800) 824-7888 (Operator 246)
In California (800) 857-7777 (Operator 246)

Back issues (Post-paid) UK, £2 00, USA & Canada, US$6 00 (surface), US$8 50 (air), Rest of World, £2 50 (surface), £3 50 (air)

***Nature* Binders**
Single Binders UK, £3 50, Rest of World, $8 50,
Set(s) of Binders UK, £9 50, Rest of World, $22 50

***Nature* Annual Indexes (1971-1980)**
Prices (post-paid)
UK, £5 00 each, Rest of World, $10 00

***Nature* First Issue Facsimile**
UK, 75p, Rest of World (surface), $1 50, (air), £2 00

***Nature* Directory of Biologicals**
(Hardcover edition) *(to be published 1981)*
UK, £18 50, Rest of World, $45 00

***Nature* Wallcharts**
How Britain Runs its Science, How the US Runs its Science, UK, £3 50, USA, $7 95 each

***Nature* in Microform**
For Information
UK UMI, 18 Bedford Row, London WC2R 4EJ
USA 300 North Zeeb Road, Ann Arbor, Mi 48106

Orders (with remittance) to
Nature, Macmillan Journals Ltd ,
Brunel Road, Basingstoke, Hants RG21 2XS, UK
Tel (0256) 29242 Telex 858493

M MACMILLAN JOURNALS

Thrombolysis by human tissue plasminogen activator and urokinase in rabbits with experimental pulmonary embolus O Matsuo, D C Rijken & D Collen 590

Demonstration of a specific dynorphin receptor in guinea pig ileum myenteric plexus C Chavkin & A Goldstein 591

Glutamate-preferring receptors regulate the release of D-[^{3}H]aspartate from rat hippocampal slices G J McBean & P J Roberts 593

Transcription of immunoglobulin heavy-chain sequences from the excluded allele P A Ponte, M Siekevitz, R C Schwartz, M L Gefter & G E Sonenshein 594

MISCELLANY

100 years ago 536

Classified advertising xvi

GUIDE TO AUTHORS

Authors should be aware of the diversity of *Nature*'s readership and should strive to be as widely understood as possible

Review articles should be accessible to the whole readership Most are commissioned, but unsolicited reviews are welcome (in which case prior consultation with the office is desirable)

Scientific articles are research reports whose conclusions are of general interest or which represent substantial advances of understanding The text should not exceed 3,000 words and six displayed items (figures plus tables) A draft abstract (50 words) is required

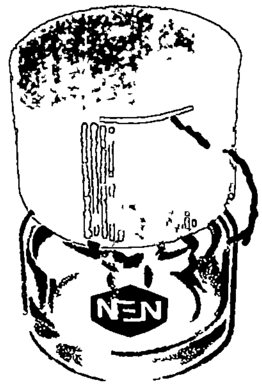
Letters to *Nature* are ordinarily 1,000 words long with no more than four displayed items The first paragraph (not exceeding 150 words) should say what the letter is about, why the study it reports was undertaken and what the conclusions are

Matters arising are brief comments (up to 500 words) on articles and letters recently published in *Nature*

Manuscripts may be submitted either to London or New York (decisions being made only in London) Manuscripts should be typed (double spacing) on one side of the paper only Three copies are required, each accompanied by copies of lettered artwork No title should exceed 80 characters in length Reference lists, figure legends, etc should be on separate sheets, all of which should be numbered Abbreviations, symbols, units, etc should be identified on one copy of the manuscript at their first appearance

References should appear sequentially indicated by superscripts in the text and should be abbreviated ~~according~~ to the *World List of Scientific Periodicals*, fourth edition (Butterworth 1963–65) The first and last page numbers of each reference should be cited References to books should clearly indicate the publisher and the date and place of publication Unpublished articles should not be formally referred to unless accepted or submitted for publication, but may be mentioned in the text

Each piece of artwork should be clearly marked with the author's name and the figure number Original artwork should be unlettered Suggestions for cover illustrations are welcome Original artwork (and one copy of the manuscript) will be returned when manuscripts cannot be published



New Carbohydrates $[^3H]$ that come clean. Vial to media in one step.

Carbohydrates $[^3H]$ in sterile aqueous solution

New convenience – no inhibiting ethanol to remove
Widest range of products

Glycoprotein Biosynthesis

Introducing at highest specific activity

Fucose,L-[6-^{3}H]-70-90Ci/mmol	NET-349X*
Galactose,D-[4,5-^{3}H(N)]- 40-60Ci/mmol	NET-698*
Acetyl-D-mannosamine,*N*- [*mannosamine*-6-^{3}H(N)]- >10Ci/mmol	NET-535A*
Fucose,L-[5,6-^{3}H(N)]-	NET-516A
Fucose,L-[6-^{3}H]-	NET-349A
Galactose,D-[1-^{3}H(N)]-	NET-126A
Glucosamine,D-[1,6-^{3}H(N)]- 30-60Ci/mmol	NET-557A*
Glucosamine,D-[6-^{3}H(N)]-	NET-190A
Glucose,D-[2-^{3}H(N)]-	NET-238A
Glucose,D-[3-^{3}H]-	NET-331A
Glucose,D-[6-^{3}H(N)]-	NET-100A
Mannose,D-[2-^{3}H(N)]-	NET-570A

Neurofunctional Mapping

Exclusive

Deoxy-2-fluoro-D-glucose,2-[5,6-^{3}H]- 25-50Ci/mmol	NET-702*
Deoxy-D-glucose,2-[^{3}H(G)]-	NET-328A
Deoxy-D-glucose,2-[1,2-^{3}H]- 30-60Ci/mmol	NET-549A*

Products designated "A" are also still available in ethanol solution

*Highest specific activity available

Circle No 02 on Reader Enquiry Card

Now dCTP $[\alpha\text{-}^{32}P]$ >7000Ci/mmol

Highest specific activity available

Highest radioactive labeling of DNA by nick translation ~2 5x10^{9}dpm/μg DNA

Fresh lots available every other week

Deoxycytidine 5′-triphosphate, tetra (triethylammonium) salt,[α-^{32}P]-
>7000Ci/mmol
10mM tricine (pH 7 6) at ~10mCi/ml, shipped in dry ice
NEG-013Z 500μCi 1mCi

Circle No 03 on Reader Enquiry Card

Latest Neuropeptides $[^{125}I]$

Adrenocorticotropic hormone,[^{125}I]- **(human)**(ACTH)
>30μCi/μg
Lyophilized from phosphate buffer (pH 3 5) containing a stabilizer and proteolytic enzyme inhibitor
NEX-165 10μCi 2x10μCi 50μCi 2x50μCi

Substance P (8-L-tyrosine),[^{125}I]-
>500μCi/μg
Lyophilized from sodium carbonate (pH 9 9) containing bovine serum albumin
NEX-152 10μCi 2x10μCi 50μCi 2x50μCi

Bradykinin (8-tyrosine)-triacetate,[8-*tyrosyl*-^{125}I]-
>500μCi/μg
Lyophilized from 0 075M ammonium acetate buffer (pH 5 0)
NEX-097 10μCi 2x10μCi 50μCi 2x50μCi

Circle No 04 on Reader Enquiry Card

Avidin-EGF-Gelatin

High biotin binding affinity
Avidin,[^{125}I]-
30-60μCi/μg
Shipped frozen in 0 05M sodium phosphate buffer (pH 7 4)
NEX-164 10μCi 2x10μCi 50μCi 2x50μCi

Tested for binding and displacement in RIA
Epidermal growth factor,[^{125}I]-
150-200μCi/μg
Lyophilized from sodium phosphate buffer (pH 7 4)
NEX-150 10μCi 2x10μCi 50μCi 2x50μCi

High fibronectin binding affinity
Gelatin,[^{125}I]- **(porcine)**
5-10μCi/μg
Lyophilized from sodium phosphate buffer (pH 7 4)
NEX-156 100μCi 500μCi

Circle No 05 on Reader Enquiry Card

Not for use in humans or clinical diagnosis

NEN New England Nuclear
®
549 Albany Street, Boston, Mass 02118
Call toll-free 800-225-1572 Telex 94-0996
(In Massachusetts and International 617-482-9595)

NEN Chemicals GmbH D-6072 Dreieich, W Germany
Postfach 401240, Tel (06103) 85034, Telex 4-17993 NEN D
NEN Canada 2453 46th Avenue Lachine, Que H8T 3C9
Tel 514-636-4971, Telex 05-821808

nature

Volume 291 No 5817 25 June 1981

Congress and cancer research 601
No Irish science politics 601
Luddites of two kinds 602

NEWS

Fusion researchers face fight for funds
Fredrickson quits without warning
European space success
Soviet space research
Telescope gets director
Endangered species
UK nuclear power
Basel Immunology Institute
German nuclear reprocessing
Water works for Cunas
European environment 607

CORRESPONDENCE

Huxley on *Nature*/Credit Nabi/etc 608

NEWS AND VIEWS

Cosmic quadrupole? (R A Muller)
Sexual differentiation of the brain (B S McEwen)
Portuguese workshop on young stars
Biogeography: in search of principles (C Patterson)
Planet X: is it necessary? (D W Hughes)
Muscle regulation: a decade of the steric blocking model (J Squire)
Exons and the evolution of haemoglobin (C C F Blake) 609

BOOK REVIEWS

Society, Technology and Risk Assessment (J Conrad, *ed*) and **Dealing with Risk (R F Griffiths, *ed*)** E Ashby, **Science & Civilisation in China, Vol V 4 (*by* J Needham)** R G W Anderson, **Neutron Spin Echo (F Mezei, *ed*)** B E F Fender, **The Foraging Strategy of Howler Monkeys (*by* K Milton)** R D Martin, **The Social Life of Monkeys and Apes, 2nd Edn (*by* S Zuckerman)** R A Hinde, **Oort and the Universe (H van Woerden *et al* , *eds*)** D Lynden-Bell, **Rheology of Polymers (*by* G V Vinogradov and A Ya Malkin)** J J Hermans, **Sociobiology Examined (A Montagu, *ed*)** J Seger 685

MISCELLANY

100 years ago 613
Product review· Scientific photography xiii

Cover

A rifle bullet piercing the skin of a soap bubble, producing V-shaped shock waves (J Kim Van Diver, MIT, Science Photo Library) The development of high-speed photography is reviewed on page 617

High-speed photography (a review) G H Lunn 617

ARTICLES

Gravitational focusing and the association of distant quasars with foreground galaxies	C R Canizares	620
Sunspot periods in the late Precambrian glacial climate and solar–planetary relations	G E Williams	624
Hormonal inducibility of rat α_{2u} globulin genes in transfected mouse cells	D T Kurtz	629
An aminoacyl tRNA synthetase binds to a specific DNA sequence and regulates its gene transcription	S D Putney & P Schimmel	632

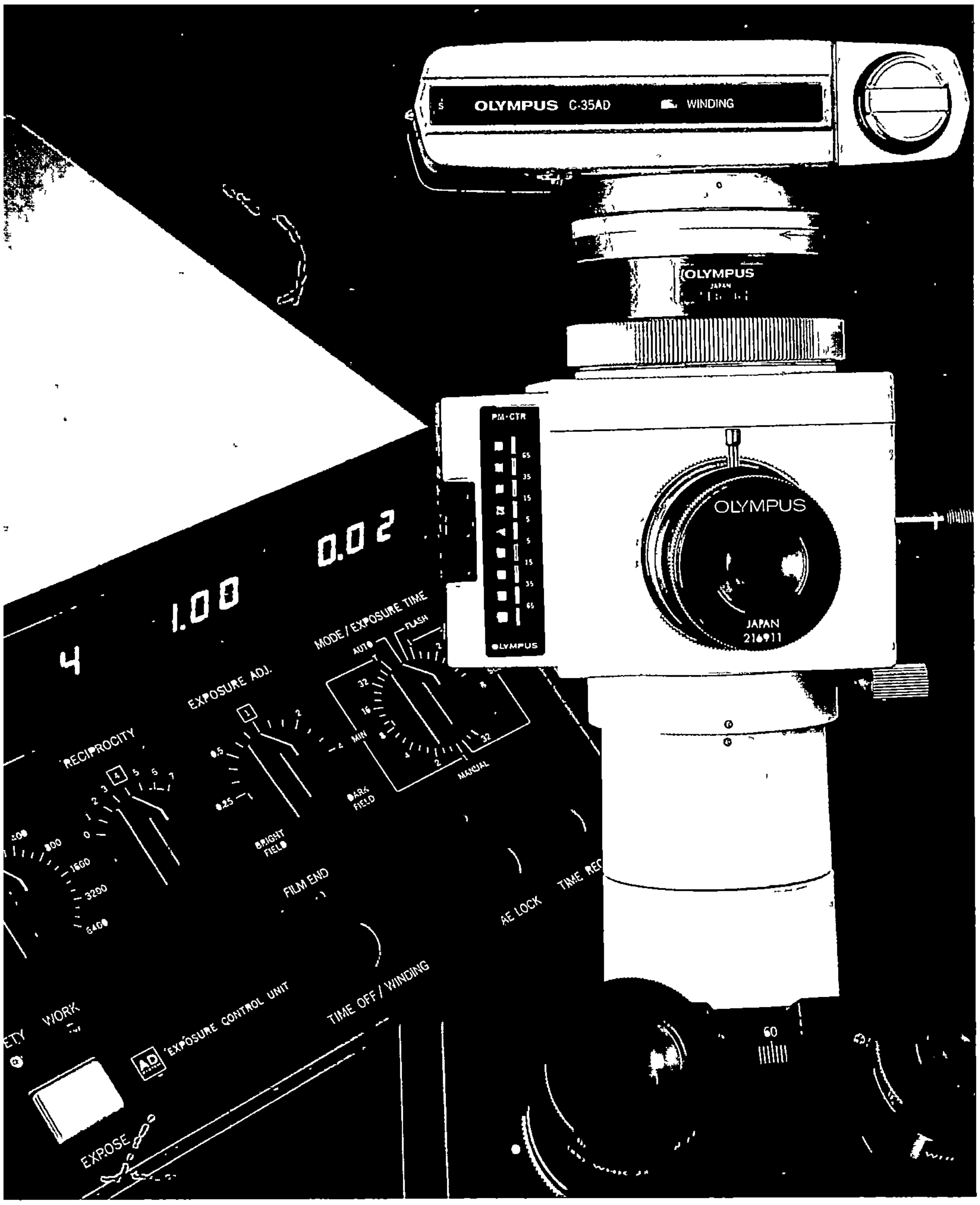
LETTERS TO NATURE

Newtonian gravity measurements impose constraints on unification theories	G W Gibbons & B F Whiting	636
Meteoric compounds in the stratosphere	A C Aikin	638
A physicochemical mechanism for the ignition of the Seveso accident	T G Theofanous	640
Proterozoic magnetic overprinting of Archaean rocks in the Canadian Superior Province	L D Schutts & D J Dunlop	642
Low-angle normal faults in the Basin and Range Province: nappe tectonics in an extending orogen	B Wernicke	645
Deep seismic soundings in India and the origin of continental crust	K R Chowdhury & R B Hargraves	648
The earliest known Palaeocene mammal fauna and its implications for the Cretaceous–Tertiary transition	J D Archibald	650
Paternity and inheritance of wealth	J Hartung	652
Sustained potentiation of transmitter release by adrenaline and dibutyryl cyclic AMP in sympathetic ganglia	K Kuba, E Kato, E Kumamoto, K Koketsu & K Hirai	654
Functional EGF receptors are present on mouse embryo tissues	E D Adamson, M J Deller & J B Warshaw	656
(—) Maackiain acetate specifically inhibits different forms of aryl hydrocarbon hydroxylase in rat and man	H V Gelboin, D West, E Gozukara, S Natori, M Nagao & T Sugimura	659
Idiosyncratic platelet responses to human tumour cells	E Bastida, A Ordinas & A Jamieson	661
Routes of conjugation in normal and cancerous tissue from human lung	G M Cohen, E M Gibby & R Mehta	662

Nature® (ISSN 0028-0836) is published weekly, except the last week in December, by Macmillan Journals Ltd Annual subscription for USA and Canada US $173 Orders (with remittance) and change of address labels to Macmillan Journals Ltd, Brunel Rd Basingstoke RG21 2XS, UK Second class postage paid at New York NY 10010 and additional mailing offices US Postmaster send form 3579 to *Nature*, 15 East 26 Street, New York, NY10010

0028-0836/81/26000C-02$01 00

Contents continued overleaf



For photomicrography or cinemicrography, it pays to match Olympus microscopes with something equally outstanding The Model PM-10AD The PM-10AD faithfully records the high image resolution, definition and clarity of Olympus microscopes And it does so with the convenience of fully automatic exposure

The time and effort you save are remarkable Thanks to a built-in microcomputer and a Silicon Blue Cell (SBC) that relays information to it quickly and accurately, you need no longer worry about complicated adjustments

The automatic exposure range is extremely wide, from 1/5000 of a second to 2 hours The microcomputer automatically corrects the reciprocity failure characteristics of the film you are using and operates continuously to yield perfect exposures

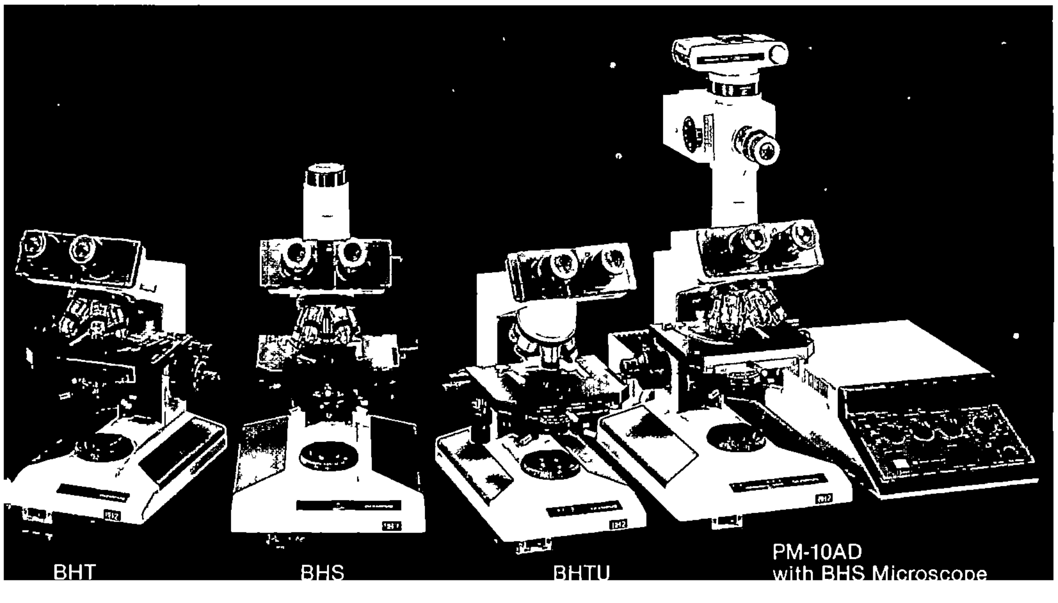
The PM-10AD allows automatic and manual operation In either mode, it can be used not only for conventional purposes but also for multiple exposures of double-stained fluorescence specimens and panorama photography, in which case uniform density is assured

Other features

- Clear digital displays include estimated and actual exposure times, time remaining to complete exposure and recall of the previous exposure time
- Unique vibration-free electromagnetic shutter The contact-free design provides a shutter life of at least 100,000 operations with absolutely no adverse effects on image quality
- Optional PM-CTR color temperature meter module for simple checking and adjustment of color temperature

A TRIUMPH OF INSPIRED SIMPLICITY.

PM-10AD Photomicrography and Cinemicrography System



BHT BHS BHTU PM-10AD with BHS Microscope

The PM-10AD is a triumph of inspired simplicity—a simplicity that means optimum results

OLYMPUS®

OLYMPUS OPTICAL CO, LTD
San Ei Building 22 2 Nishi Shinjuku 1 chome Shinjuku ku Tokyo Japan
OLYMPUS OPTICAL CO (EUROPA) GMBH
Postfach 104908 Wendenstrasse 14 16 2 Hamburg 1 West Germany
OLYMPUS CORPORATION OF AMERICA
4 Nevada Drive New Hyde Park N Y 11042 U S A

Circle No 19 on Reader Enquiry Card

• **How Britain runs its Science**

• **How the US runs its Science**

These two new *Nature* wallcharts are now available.

The 2 x 3ft guides offer a comprehensive survey of the administration of science in Britain and the US, showing the interlocking relationship between science, engineering and medical establishments

Designed as fold-out sheets, the wallcharts can be carried easily to meetings or hung as posters for permanent reference

The new 1981 wallcharts are available at £3 50 or $7 95 each from *Nature's* offices in London, Basingstoke and New York

Annual Subscription including Index (51 issues)

UK & Eire		£75
USA & Canada	Surface only	US$173
Belgium	Airspeed only	BF5695
West German	Airspeed only	DM360
Netherland	Airspeed only	G390
Switzerland	Airspeed only	SF330
Rest of Europe	Airspeed only	£85
Rest of World	Surface	£85
Rest of World*	Airmail	£130

**(not USA, Canada & Europe)*

Personal subscription rates are available in some countries to subscribers paying by personal cheque or credit card

Details can be obtained from
UK *Nature* Promotion Department,
4 Little Essex Street, London WC2R 3LF
Telephone Jonathan Earl 01-836 6633

USA *Nature*, 15 East 26 Street, New York, NY 10010
Credit card orders only (in USA & Canada)
Call toll-free (800) 824-7888 (Operator 246)
In California (800) 857-7777 (Operator 246)

Back issues (Post-paid) UK, £2 00, USA & Canada, US$6 00 (surface), US$8 50 (air), Rest of World, £2 50 (surface), £3 50 (air)

***Nature* Binders**
Single Binders UK, £3 50, Rest of World, $8 50,
Set(s) of Binders UK, £9 50, Rest of World, $22 50

***Nature* Annual Indexes (1971-1980)**
Prices (post-paid)
UK, £5 00 each, Rest of World, $10 00

***Nature* First Issue Facsimile**
UK, 75p, Rest of World (surface), $1 50, (air), £2 00

***Nature* Directory of Biologicals**
(Hardcover edition) *(to be published 1981)*
UK, £18 50, Rest of World, $45 00

***Nature* Wallcharts**
How Britain Runs its Science, How the US Runs its Science, UK, £3 50, USA, $7 95 each

***Nature* in Microform**
For Information
UK UMI, 18 Bedford Row, London WC2R 4EJ
USA 300 North Zeeb Road, Ann Arbor, Mi 48106

Orders (with remittance) to
Nature, Macmillan Journals Ltd ,
Brunel Road, Basingstoke, Hants RG21 2XS, UK
Tel (0256) 29242 Telex 858493

M MACMILLAN JOURNALS

Sequential appearance of Epstein–Barr virus nuclear and lymphocyte-detected membrane antigens in B cell transformation D J Moss, A B Rickinson, L E Wallace & M A Epstein 664

Inserted H-2 gene membrane products mediate immune response phenotype of antigen-presenting cell A Jakobovits, A Frenkel, N Sharon & I R Cohen 666

Differences between germ-line and rearranged immunoglobulin V_x coding sequences suggest a localized mutation mechanism M Pech, J Hochtl, H Schnell & H G Zachau 668

Bleomycin-resistant DNA synthesis in ataxia telangiectasia cells P Cramer & R B Painter 671

Absence of luteinizing hormone-releasing and anti-fertility properties in a glucagon-selective somatostatin analogue F J Bex, A Corbin, D Sarantakis & E Lien 672

Multiple sequences related to classical histocompatibility antigens in the mouse genome B Cami, F Brégégère, J P Abastado & P Kourilsky 673

Homologous tyrosine phosphorylation sites in transformation-specific gene products of distinct avian sarcoma viruses J C Neil, J Ghysdael, P K Vogt & J E Smart 675

The structure of a chromosomal leghaemoglobin gene from soybean E Ø Jensen, K Paludan, J J Hyldig-Nielsen, P Jørgensen & K A Marcker 677

A human proinsulin variant at arginine 65 D C Robbins, P M Blix, A H Rubenstein, Y Kanazawa, K Kosaka & H S Tager 679

Allosteric regulation of crocodilian haemoglobin M F Perutz, C Bauer, G Gros, F Leclercq, C Vandecasserie, A G Schnek, G Braunitzer, A E Friday & K A Joysey 682

GUIDE TO AUTHORS

Authors should be aware of the diversity of *Nature*'s readership and should strive to be as widely understood as possible

Review articles should be accessible to the whole readership Most are commissioned, but unsolicited reviews are welcome (in which case prior consultation with the office is desirable)

Scientific articles are research reports whose conclusions are of general interest or which represent substantial advances of understanding The text should not exceed 3,000 words and six displayed items (figures plus tables) A draft abstract (50 words) is required

Letters to *Nature* are ordinarily 1,000 words long with no more than four displayed items The first paragraph (not exceeding 150 words) should say what the letter is about, why the study it reports was undertaken and what the conclusions are

Matters arising are brief comments (up to 500 words) on articles and letters recently published in *Nature*

Manuscripts may be submitted either to London or New York (decisions being made only in London) Manuscripts should be typed (double spacing) on one side of the paper only Three copies are required, each accompanied by copies of lettered artwork No title should exceed 80 characters in length Reference lists, figure legends, etc should be on separate sheets, all of which should be numbered Abbreviations, symbols, units, etc should be identified on one copy of the manuscript at their first appearance

References should appear sequentially indicated by superscripts in the text and should be abbreviated according to the *World List of Scientific Periodicals*, fourth edition (Butterworth 1963–65) The first and last page numbers of each reference should be cited References to books should clearly indicate the publisher and the date and place of publication Unpublished articles should not be formally referred to unless accepted or submitted for publication, but may be mentioned in the text

Each piece of artwork should be clearly marked with the author's name and the figure number Original artwork should be unlettered Suggestions for cover illustrations are welcome Original artwork (and one copy of the manuscript) will be returned when manuscripts cannot be published

SCIENTIFIC & MEDICAL DIVISION
MACMILLAN PUBLISHERS

THE ECOLOGY & CONSERVATION OF LARGE AFRICAN MAMMALS

S. K. ELTRINGHAM

DEPARTMENT OF APPLIED BIOLOGY,
UNIVERSITY OF CAMBRIDGE.

£15.00; 288pp; ISBN 0 333 23580 0.

"Keith Eltringham has produced a valuable book which treats the ecology of large African mammals in the widest context In a compact volume he considers their social behaviour, population, biology, and feeding in an ecological framework. Perhaps the most important sections are the last three chapters which deal with conservation, wild life management, and man and national parks"
BIOLOGIST, APRIL 1980

"Eltringham's book is written as a textbook for undergraduates but he also had in mind the value of making the book of interest to those with a more cusory contact with African habitats. Accordingly, it is written with a pleasant directness, lack of jargon, and wry humour"
TIMES HIGHER EDUCATION SUPPLEMENT, MARCH 1980

"The appearance of this book is timely Many talk about conservation without doing anything about it Keith Eltringham has done an excellent job by providing much needed information. He has obviously drawn from his vast experience as a wildlife biologist in East Africa for many years Writing about a subject he knows so well makes the account vivid and very much alive"
AFRICAN JOURNAL OF ECOLOGY, MARCH 1980.

"This immensely readable book will be essential for any undergraduate studying mammals and ecology but it will also be attractive to a wider audience – not only those who have been fortunate enough to visit Africa Many people are now familiar with these animals through films, television, zoos and safari parks, and this book will interest all those who wish to know how they are living together and surviving in the world
BRITISH BOOK NEWS, APRIL 1980

Macmillan books are available from all leading booksellers.

In case of difficulty write to Frances Roach, Macmillan Retail Ltd., Houndmills, Basingstoke, Hampshire, England, enclosing remittance plus 70 pence (p&p).

Circle No 50 on Reader Enquiry Card

©1981 Macmillan Journals Ltd
Nature® is published weekly
ISSN 0028-0836
Registered as a newspaper at the British Post Office

EDITORIAL OFFICES

London
4 Little Essex Street, WC2R 3LF
Telephone (01) 836 6633 Telex 262024
Telegrams Phusis London WC2R 3LF

Editor John Maddox
Deputy Editor Peter Newmark
Editorial Staff
Alun Anderson Sara Nash
Philip Campbell Peta Pickering
Isobel Collins Judy Redfearn
Konrad Guettler Miranda Robertson
Tim Lincoln Robert Walgate
Naomi Molson Charles Wenz

Washington News Bureau
801 National Press Building, DC 20045
Telephone (202) 737-2355 Telex 64280
David Dickson (Washington News Editor)

Publisher Elizabeth Hughes
Marketing Director Ray Barker
International Advertising Manager
Andy Sutherland
Features Advertising Manager Marion Delaney
Promotion Manager Jonathan Earl

New York
15 East 26 Street, New York, NY 10010
Telephone (212) 689-5900
American Publisher Robert Ubell
American Advertising Manager Henry Dale
Marketing Manager Sheila Kane

Classified Advertising
Jean Neville, 4 Little Essex Street,
London WC2R 3LF
Telephone (01) 240 1101

US Classified Advertising
Cathy Moore, 15 East 26 Street,
New York, NY 10010
Telephone (212) 689-5900

Canadian Display and Classified
Peter Drake, 32 Front Street West,
201 Toronto, Ontario M5J IC5
(416) 364-1623

US Display Advertising Representatives
Nature New York Office—(212) 689-5900
Roy McDonald Associates, Inc
(Dallas)—(214) 941-4461
Roy McDonald Associates, Inc
(Houston)—(713) 988-5005
Didier & Broderick, Inc (Chicago)—
(312) 498-4520
Jobson/Jordan/Harrison/Schulz, Inc
(San Francisco) (415) 392-6794
(Los Angeles & Pasadena) (213) 796-9200
Daniel Adams Associates, Inc
(Philadelphia)—(215) 353-6191
CEL Associates (Boston)—(617) 848-9306
Brinker & Brinker (Fort Lauderdale)
(305) 771-0064

Japanese Advertising Representative
Orient Echo Inc
11th Fl, Sankyo Bldg
7-12, Tsukiji 2-Chome,
Chuo-Ku, Tokyo, Japan
Telephone 541-4923

Vol. 291 No. 5817 25 June 1981

nature

7 May 1981

Reagan's need of a proliferation policy

Another bout of anxiety about the proliferation of nuclear weapons is plainly in prospect. The announcement last week by the government of South Africa that it had started to manufacture enriched uranium is certain to rekindle anxiety about the republic's long-term intentions. Driblets of news, not easily verified, from Pakistan, seem calculated to have the same effect. And the disagreement between the United States and India about the supply of American enriched uranium for Indian reactors seems finally to have led to the proposal that the treaty between the two countries on the supply of nuclear materials, first signed in 1963, should now be abandoned. The consequence will be that the Indian Atomic Energy Commission will be able to use the reactors covered by the agreement as it wishes, and that there will be no external restraints (except the shortage of resources) on its enrichment of uranium. Taken separately, these developments are not as worrying as they may seem to some. A more substantial cause for anxiety is that they coincide with the Reagan Administration's failure to work out its policy on nuclear proliferation.

One difficulty, not as far beyond President Reagan's control as it may seem in what has been called his second honeymoon, is that the nomination of Mr James Malone as the assistant secretary at the State Department responsible for the non-proliferation policy is held up in the Senate. That, however, need not mean that nothing can be decided. The most urgent need is that the new Administration, which cannot welcome being bound (as it still is) by the terms of President Carter's Anti-Proliferation Act, should decide whether it wishes to amend the act, and how, or whether it should simply repeal it. The essential folly of the act is that it seeks to apply to all potential purchasers of nuclear materials from the United States a unilateral policy which overrides pre-existing agreements (as in the case of India), even with those states that have ratified the non-proliferation treaty (which India has not). The act is also closely linked with President Carter's interdiction against the reprocessing of nuclear fuel for the extraction of plutonium. Applying this rule to all buyers of United States uranium was scarcely palatable when the United States itself was not reprocessing spent uranium. This policy will be less defensible if, as seems inevitable, reprocessing of fuel begins again in the United States in the next few years. But even deciding what should be done domestically is hampered by the Administration's ill-preparedness. Thus the leaderless Nuclear Regulatory Commission is at sixed and sevens (see *Nature* 30 April, p 727), and unlikely to be able to contribute much to constructive policy formation.

Another imponderable is that nobody can tell how Congress would respond to suggestions that present policies should be modified. President Carter's men, when persuading Congress to enact the present legislation, did much (and more than they should have done) to persuade anybody who would listen that the risk of the emergence of nuclear weapons elsewhere must be simply proportional to the quantity of plutonium in existence. That message, however over-simple, may have stuck. Certainly Congress is unlikely willingly to give up the powers that act has almost fortuitously presented it with. Having wished the Non-Proliferation Act on Congress and himself, President Carter had to suffer the humiliation of asking Congress to make an exception of the provision of enriched uranium required of the United States by the 1963 bilateral agreement between itself and India. President Reagan cannot risk asking Congress to embark on a piecemeal amendment of existing legislation. He needs both a coherent domestic policy on the development of nuclear power, with our without reprocessing, and an external policy that makes some concessions to the principle that countries overseas need to know in advance what is required of them when they buy American nuclear materials. And, even then, Congress will have to be re-educated. That, presumably, is one of Mr Malone's intended tasks.

So what should the Reagan Administration require of its nuclear customers overseas? The Non-Proliferation Act requires that potential customers should sign the Non-Proliferation Treaty (which previously unwilling Egypt has now done). In that respect, the United States is asking more of its customers than other suppliers of nuclear materials, France for example. There is, however, nothing wrong with that provision — would that other suppliers made the same demands. The iniquity of the act is that even those customers who have ratified the treaty, and thus bound themselves both not to carry out nuclear explosions and to open all nuclear installations to the international inspectors, are required not to extract plutonium from their spent fuel. (Technically, the act allows reprocessing if the objective is to obtain plutonium for research on fast reactors, but only then with the prior agreement of the United States.) President Carter appeared not to have appreciated how offensive were these provisions overseas. Some states can fairly argue that they need to reprocess spent uranium for economic reasons. All non-nuclear powers among them can rightly say that if they have, at the urging of the United States, ratified the Non-Proliferation Treaty, it is inequitable that further restrictions should then be arbitrarily imposed. Can it be, they say, that the United States does not believe in the international system of safeguards that it has helped to create? Worse still, does the United States assume that states bound by international treaties will nevertheless break their word whenever they have the chance?

The plain truth, of course, is that there is no necessary connection between the amount of plutonium within a country's borders and the chance that it will go on to make nuclear weapons. There is plenty of evidence that the factors that drive states to one policy or another are political, not technical. Except for the five well-advertised nuclear powers, political considerations appear so far to have argued for restraint and propriety. Israel has gone further along the nuclear road than most other non-nuclear states, and has been compelled to do so in a significant way. The word has been put about that nuclear technology in Israel is advanced, and that there may be plutonium sufficient for several weapons tucked secretly away. So Israel is a kind of crypto-nuclear power. The snag (from the Israeli point of view) is that it would be exceedingly dangerous to carry out a nuclear test, or openly to confess that there is a stock of bombs. For then the wrath not merely of the Middle East but of the super-powers would be visited on the Israeli government. It is not unreasonable to guess that South Africa may be dickering with a similar policy. The Indian nuclear explosion in 1974 (which prompted the still unanswered question of when the second explosion will take place) had more ambiguous motives, but served quickly to remind the Indian government of its dependence on economic assistance from overseas.

The Reagan Administration is as well placed as its predecessors to manipulate political pressures such as these, and thus to inhibit the proliferation of nuclear weapons. Domestic politics would probably prevent the Administration from demanding that Israel should sign the Non-Proliferation Treaty (now that Egypt has

0028-0836/81/190001-02$01 00

done so) if it is to continue receiving assistance with conventional armaments Pakistan would be an easier target, however — and Pakistan's signature of the treaty would go a long way to persuade India that playing about with bombs is not worthwhile The strangest feature of President Carter's advocacy of the Non-Proliferation Act of 1978 is that, although used to decrying what are called "technological fixes" in other contexts, he relied on just such a device in the battle against proliferation. The question now is whether President Reagan will be more subtle He will have to persuade Congress that the American influence on the pace of proliferation is best left to him and to the State Department, not left to the wooden interpretation of an over-restrictive piece of legislation At the same time, he will have to persuade both Congress and its electors that nuclear power is not necessarily the end of the world The more time that slips by, the harder will be the task

Advice and dissent on civil science

The inquiry by the House of Lords Select Committee on Science and Technology seems already to have unearthed a substantial problem, the constitution of the committee known as the Advisory Board on the Research Councils (see page 3) Much of the evidence so far sent in to the committee, that from the Royal Society for example, has argued that the board's interpretation of its function has been unhelpfully narrow Constitutionally the board is advisory to the Secretary of State for Education and Science, who is in turn responsible for writing the cheques that keep the research councils alive Since it was set up a decade ago as part of the Rothschild reorganization of civil science, the board has been almost exclusively a forum for horse-trading between the research councils. Which budgets should be increased more quickly than inflation, which should be trimmed? Given the inevitable dominance of the board by the powerful figures of the chief executives of the research councils, and the prospect that research budgets will indefinitely stagnate, the board's concentration on mere money and its distribution is understandable. The complaint is that, in the circumstances, there is no mechanism for tackling more general questions central to the administration of civil science

The charge is valid even though the board seems recently (to its credit) to have taken on some general questions of common concern to the research councils as a group Thus the board has begun an investigation of the relationship in the funding of research of the University Grants Committee and the research councils — and has made very little progress There is an inquiry into graduate education and the basis on which research councils support graduate students in universities and other institutions, and also into the employment (and fate) of postdoctoral researchers not members of academic staffs In due course, if the busy people who have committed themselves to work on these questions can find the time, the research councils may be helped to a better understanding of some important questions. But they are largely domestic questions They do not apparently include the long overdue appraisal of the Rothschild reorganization which among other things gave operational departments charge of part of the budget for civil science Instead, the board has lent its name to the piecemeal dismantling of the Rothschild system typified last year by the transfer back to the Medical Research Council of funds earlier transferred to the Department of Health

The vein of discontent with the advisory board that the House of Lords has tapped is in no sense a new discovery Ironically, however, the narrowness of the board's terms of reference is itself a consequence of the Rothschild reorganization In 1971, the Council for Scientific Policy (the board's predecessor in the 1960s) and Lord Rothschild independently embarked on a consideration of how the British government's support for civil science should be organized The council concluded that there should be an advisory committee (called a board) with wide powers to consider the research policies of government departments and research councils. Lord Rothschild eventually came to the opposite conclusion — that the Council for Scientific Policy should become strictly advisory, and that its chief task should be to recommend the division of the funds available between the research councils Sir Frederick Dainton, the council's chairman at the time and already incensed that his own recommendations were kept secret while the Rothschild inquiry was under way, was understandably infuriated that the government of the day should have accepted Rothschild's view that the committee's terms of reference should be narrow but should then have called that watered-down version of the standing advisory committee a "board", almost the only proposal of Dainton's to be accepted. The implication is that the advisory board, however unsatisfactory, has been doing this past decade what was expected of it.

The temptation now, in the House of Lords and elsewhere, will be to tinker with terms of reference. Sir Alec Merrison, the chairman for the time being of the board (and of a host of other committees) said last week that his board should have a wider remit but should not be encouraged to poach on the ground occupied by another advisory committee, the Advisory Council on Applied Research and Development. The modesty is seemly but inappropriate. There is no reason why broader terms of reference should of themselves allow the committee to function more effectively. If it cannot get through the work it has already set itself, what reason is there to expect the same committee with a similar way of working to cover more ground effectively? At the same time, now that the board has come into being, and seems anxious (given time and resources) to tackle a variety of problems common to the research councils, it would be a misfortune if yet another reorganization should bring this good work to a stop

Where, then, should the House of Lords look for a solution? The most glaring weakness of the advisory board is its lack of independence from the research councils as such (which is not to disparage the work of the independent members of the board) The advisory council, theoretically more influential because it reports to the Cabinet Office rather than to a department, and whose members are all independent persons, suffers from being if anything too detached from the problems that concern government from day to day. The result is, so far as the public knows, that the council has spent its time drawing attention to fields of science or technology in which research might be expected to be worthwhile. Structural questions — the relationship between the research councils and the University Grants Committee for example — seem not to have arisen Yet these are precisely the questions on which the British government needs not merely advice but policies. The obvious solution is to leave the Advisory Board for the Research Councils much as it is, with responsibility for advice on the coordination of largely academic science, but to strengthen the advisory council For now that the research councils have begun talking to each other in the presence of outsiders about the coordination of their work, it would be a pity ot throw away the opportunities thus provided Strengthening the advisory council is more than merely a matter of redrafting terms of reference. With the best will in the world (and even with a chairman as able as Sir Alfred Spinks), busy people meeting occasionally cannot be expected to do the hard work that circumstances require At the least, the chairman should be a full-time appointment (in which case he or she might usefully double as chairman of the advisory board). But the committee also needs a professional staff, not mere minute-takers represented by promising young recruits to the civil service Both changes would be easier if, as in the 1950s, responsibility for civil science were removed from the Department of Education and Science. They, and the others in prospect, would have more meaning if some British government would believe its own declaration about the importance of this subject.

0028-0836/81/190002 01$01 00

Recombinant DNA guidelines wear thin

NIH moves to voluntary regulation

Washington

The US National Institutes of Health (NIH) have taken the first steps towards removing mandatory controls on the conduct of recombinant DNA research. Meeting in Bethesda, Maryland, the institutes' Recombinant DNA Advisory Commitee (RAC) set up a subcommittee to discuss whether the present guidelines, which all federally-sponsored scientists have to observe, should be turned into a voluntary code of practice, and the answer is likely to be yes.

The committee's decision was in response to a proposal from two of its members, Dr Allan Campbell of Stanford University and Dr David Baltimore of the Massachusetts Institute of Technology. They suggested that the guidelines should be voluntary — eliminating the formal need for review bodies such as local Institutional Biosafety Committees — and that there should be major reductions in the containment levels recommended for various types of experiments (*Nature* 26 March, p.281).

Their proposal, which was published in the *Federal Register* on 20 March and would involve eliminating virtually all the regulatory requirements contained in Part III of the current guidelines, has generated considerable support in the scientific community.

Most members of the committee seemed to agree with the spirit of the proposition. Their willingness to relax the guidelines was illustrated in an earlier vote on reducing containment requirements for experiments involving *Escherichia coli* K12 and laboratory strain *Saccharomyces cerevisiae* host vector systems. The committee adopted the most liberal option, recommending that such experiments, making up about 80 per cent of all recombinant DNA experiments, be totally exempted from the guidelines.

Several committee members pointed out, however, that there is still considerable public concern about the potential social effects of recombinant DNA research. Hence should such research be regulated, it must be seen to be done in a responsible manner as otherwise federal or state legislation might be provoked. Other members, particularly those from the public interest community, argued against reducing the containment conditions too quickly, pointing out that although a few experiments seemed to confirm that recombinant DNA experiments posed no greater risks than other experiments with the same organisms, this hypothesis had not yet been rigorously tested. Even Dr Campbell and Dr Baltimore agreed that experiments over 10 litres should still be subject to special review by the director of NIH, since large-scale experiments might still result in unsuspected hazards.

The committee members finally agreed that, rather than recommending to NIH director Dr Donald Fredrickson the immediate transmutation of the guidelines from formal restrictions into a voluntary code of practice, a subcommittee should look into the whole question of the future of the guidelines.

The proposal made by Dr Campbell and Dr Baltimore would result in the recommendation that most experiments involving recombinant DNA be carried out in P1 containment conditions, these representing prudent laboratory practice. Potentially hazardous microorganisms would be classified and handled according to procedures recommended by the Center for Disease Control (CDC) in Atlanta, Georgia. Several scientists have expressed concern about proposed revisions to these procedures, published by CDC last year, arguing that in certain cases they would be more restrictive than necessary. However, Dr Baltimore told the committee that he believed CDC "will eventually produce a document that is a useful guide to the appropriate laboratory practice".

Dr Baltimore also stressed that RAC would still be needed to maintain surveillance over recombinant DNA research in the future, to make decisions about experiments which may present particular dangers and to help in the evaluation of ethical issues, which were going to become a very important part of a recombinant DNA activities.

David Dickson

Plea to keep guidelines

A report published by the Office of Technology Assessment, after a two-year study of the impact of applied genetics, says that there are several reasons for not eliminating the guidelines completely, pointing out that "sufficient scientific concern exists for the guidelines to prohibit certain experiments and to retain containment for others", and that most experiments can already be carried out at the least burdensome containment levels.

Although conceding that NIH showed flexibility in liberalizing the restrictions as evidence showed lower risk levels than originally feared, there is criticism of the way that NIH have played the role of oversight body. Several shortcomings are listed, including the virtual non-existence of monitoring for compliance to the guidelines, and the lack of systematic evaluation of other techniques, such as cell fusion, that might present risks.

The report, which lists some of the many applications of recombinant DNA techniques in a wide range of industries, says that Congress might wish to consider the need for regulating work with all hazardous microorganisms and viruses, whether or not they are products of genetic engineering.

Role of advisory committee questioned

The United Kingdom's Advisory Board for the Research Councils could provide government with advice on a wide range of public policy issues with a scientific content if its terms of reference were extended, according to its chairman, Sir Alec Merrison. At present, the board is constrained by its remit to advise the Secretary of State for Education and Science chiefly on the distribution of the science vote amongst the research councils.

The board is like a "huge torpedo" that is not being launched "at anything very much" Sir Alec told the House of Lords Select Committee on Science and Technology last month as part of its investigation into the provision of scientific advice to government. The expertise of the board's members, senior scientists and administrators, could be put to better use if it took on a role similar to the Advisory Council on Scientific Policy, which provided advice on science policy through a chief scientist to the government before its demise in 1963. Any extension of the board's remit, however, would require very careful thought, said Sir Alec, to avoid treading on the toes of the Advisory Council for Applied Research and Development which does a good job in advising on scientific aspects of industrial policy.

Informal contact with the Department of Education and Science was generally good, Sir Alec said. But it had not worked in the case of the government's White Paper on public expenditure, published on budget day last March. Although the implications of the White Paper for the support of basic science in the universities were "profound", the board only learnt of the government's plans in a brief letter a few days before the budget statement.

Sir Alec said that by thinking that it could protect science by maintaining the level of the science vote while considerably reducing the grant for student fees, the government had clearly not understood the way in which university science is funded. The implications for that system, known as the dual support system, are so grave, said Sir Alec, that work on a report that the board is preparing into its workings has had to stop.

Judy Redfearn

0028-0836/81/190003-01$01.00

Carter's science team

Yesterday's men

Washington

As the Reagan Administration continues to have difficulty in filling many of its top jobs — posts which remain vacant include the directorship of the Office of Science and Technology Policy (OSTP), the chairmanship of the Nuclear Regulatory Commission, and the directorship of the Office of Energy Research — science policy officials from the Carter Administration are having mixed success in finding new positions. Most have slipped back into academic posts — the traditional fall-back for Democrats, in contrast with Republicans, who usually return to the private sector after leaving office. But for some, the prospects are still uncertain.

Most successful in finding a new job has been Dr Frank Press, Mr Carter's science adviser and director of OSTP, who was recently elected to succeed Dr Philip Handler as president of the National Academy of Sciences. Dr Press takes up his new position in Washington at the beginning of July, and has meanwhile returned temporarily to his old post as professor of geology at the Massachusetts Institute of Technology.

Of Dr Press's three original associate directors at OSTP, Dr Gilbert Omenn, responsible for health sciences, has returned to the University of Washington in Seattle, where he is professor of medical genetics, having spent the last year of the Carter Administration as associate director of the Office of Management and Budget. Dr Philip Smith, associate director for natural resources, has temporarily joined the staff of the National Science Board, the overseeing body of the National Science Foundation. The third, Dr Benjamin Hubermann, remains at OSTP keeping the director's seat warm until a successor to Dr Press is appointed.

Dr Robert Frosch, administrator of the National Aeronautics and Space Administration (NASA) under Mr Carter, is already established as president of the newly formed American Association of Engineering Societies. Dr Hans Mark, previously secretary of the Air Force and before that director of NASA's Ames Research Laboratory, was at one point tipped to succeed Dr Frosch, but has now been nominated as deputy to NASA's new administrator, Dr James Beggs.

Another academic who has returned to her former base is Dr Eula Bingham, previously Assistant Secretary of Labor responsible for the Occupational Safety and Health Administration (OSHA). Dr Bingham has rejoined the University of Cincinnati, from which she had been on leave as professor of environmental health.

In contrast, Dr Bingham's counterpart at the Environmental Protection Agency, ex-administrator Mr Doug Costle, is entering the academic world for the first time. Before his appointment under Mr Carter, Mr Costle had been an assistant director of the Congressional Budget Office, and commissioner of environmental protection in the state of Connecticut. A lawyer by profession, he was recently named as a visiting scholar at Harvard University's School of Public Health, and a lecturer at the John F. Kennedy School of Government.

Moving in a different direction is Dr Edward Frieman, brought in to succeed Dr John Deutch as head of the Office of Energy Research. Before coming to Washington, Dr Frieman was deputy director of the Plasma Physics Laboratory at Princeton. He has now joined a Washington-based consulting firm, Science Applications Incorporated, which deals in a range of energy issues.

Two appointments have remained unchanged. Dr Donald Fredrickson, orginally appointed by President Ford as director of the National Institutes of Health for a six-year term, has stayed in that position. Also staying put is Dr John Slaughter, brought in during the last months of the Carter Administration as director of the National Science Foundation partly on the understanding that Mr Reagan would not send him straight back to Washington State University, where he had only recently been appointed provost.

Less fortunate has been Dr Anthony Robbins, previously director of the National Institute for Occupational Safety and Health (NIOSH). Dr Robbins became the target of a bitter attack by conservative groups for his support of labour union positions on some occupational health issues, and was summarily removed from his post by the new Health Secretary Mr Richard Schweiker, as well as being stripped of a commission in the US Health Corps.

At the State Department, Thomas Pickering, previously head of the Bureau of Oceans and International Environmental and Scientific Affairs, has returned to his position as a career foreign service official. And William J. Perry, previously Under-Secretary for Defense for Research and Engineering, has joined a San Francisco investment banking firm.

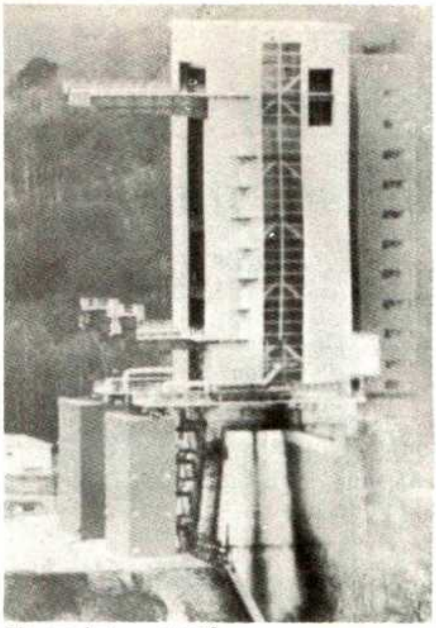
David Dickson

Ariane development

Launch at last?

The European Space Agency (ESA) has announced that the third test launch of the Ariane launcher, with two satellites on board, is to take place on 19 June. After extensive testing and some modifications, the problems associated with the first-stage engines have apparently been solved.

Europe's three-stage launcher programme encountered a major setback during the second Ariane test launch on 23 May 1980. Subsequent analysis revealed that instabilities in the injection system caused pressure oscillations at 2,300 Hz to develop inside the combustion chamber of one of the four first-stage Viking engines. The oscillations caused structural damage which in turn hindered cooling and resulted in rupture of the engine. Subsequent ground tests also revealed oscillations at 2,700 Hz.



Up, up, but . . . away?

Liquid-fuelled rocket motors are notorious for such instabilities, which still confound theorists. ESA's trial and error approach to the problem has involved months of testing of alternative modifications to the injector system. One design, in which the injection nozzle diameters have been slightly widened, resulted in a complete absence of pressure oscillations during ground tests, and has now been adopted. As a further safety measure, manufacturing tolerances have been tightened.

Like its predecessors, the third Ariane is to be launched from Kourou in French Guyana. It will carry Meteosat 2, a European meteorological satellite, and APPLE, an Indian telecommunications satellite. If all goes well, the fourth and final test launch will take place in October.

Ariane's problems have hindered its progress as competitor to the United States shuttle. However, contractors such as Intelsat — the international telecommunications satellite organization — have not cancelled their options on Ariane as a result of the delays.

Philip Campbell

British aerospace

Limbering up

Further evidence of the UK government's sudden eagerness to enter the space race came last week in the form of a new committee set up as a forum for the views of the space industry, users and government. This Space Consultative Committee has held its first meeting to try to determine its role.

The diverse membership of the committee, which includes representatives of government departments, the Science

0028-0836/81/190004-02$01.00

Research Council, the broadcasting authorities, British Telecom and satellite and telecommunications equipment manufacturers, suggests that it is intended to coordinate views from a variety of sources, a role that in other countries might be taken by a national space agency.

Mr Keith Marshall, Under-Secretary of State in the Department of Industry, who will chair the committee, said in parliament that it would also be investigating topics such as the future programmes of the European Space Agency in telecommunications and remote sensing satellites, a satellite communications system for the Ministry of Defence and the marketing strategy of British satellite and ground-station technology abroad. He also indicated that the government would expect industry to put up all the finance for promising projects.

The committee will shortly have before it the conclusions of two reports commissioned by the government. The report of the Home Office study on direct broadcasting should be published later this month and the conclusions of the report of the Central Policy Review Staff on Britain's future in space are to be made available to the committee.

Hopes that an abbreviated version of the review body's report would be published, however, were quashed at the end of last month by a statement in the House of Commons — the government apparently fears breaching industrial confidentiality by allowing open discussion of the report. So far the only conclusions to emerge have been rather bland, such as that the United Kingdom should continue its commitment to the European Space Agency. It is to be hoped that the committee will provide a focus for wider debate. **Judy Redfearn**

European Parliament

Nuclear ban debate

Brussels

A moratorium on nuclear power in Europe will be the subject of a debate in the European Parliament this week. Also on the agenda is a motion on the Geneva Appeal which was launched as a protest against France's fast breeder reactor, Super-Phénix.

The moratorium debate will focus on a report by Sir Peter Vanneck (European Democrat) who is strongly opposed to the idea. The report was adopted by the energy committee by 13 votes to 7 following a resolution by three independent Members of the European Parliament in December 1979. The moratorium would be used as an opportunity to look into the problem of storing nuclear waste, harmonizing safety standards and promoting widespread public discussion. In his report, Vanneck concedes the benefit of these points but feels that the cure would be more painful than the disease.

Depending on how widespread the moratorium would be, Vanneck calculates that by 1990 it would leave the Community short of between 45 and 115 GW, equivalent to 76 or 170 million tonnes of oil a year. By 1990, 34 per cent of the EEC's electricity needs will be supplied by nuclear power. Vanneck also points out that in 1975 electricity from nuclear power caused no deaths in the United Kingdom, while maritime transport (of oil and coal) caused 86 deaths per 100,000 employees and coal mining caused 23.4 deaths per 100,000.

The underlying cause of public concern, says Vanneck, is that plutonium produced in nuclear power stations may be used for military needs. This issue will certainly be linked with the debate on the Geneva Appeal. In October 1978, more than 30,000 French and Swiss, including a number of leading figures, signed an appeal asking the political community in Europe at large to consider alternatives to fast breeder reactors. The appeal reflects the fears of both Swiss and French living in the neighbourhood of the 1,250 MW Super-Phénix fast breeder reactor under construction at Greys Malville in France.

The Council of Europe had held public hearings on breeder reactors in 1979. The present motion, if adopted, would call on all governments concerned to suspend work connected with breeder reactor projects. The EEC itself does not directly contribute to any fast breeder construction programme. The European Commission, however, is preparing a report on the state of the nuclear fuel reprocessing industry.

Both debates, are nevertheless rearguard actions of the anti-nuclear group in the parliament. Most parliamentarians, as previous resolutions have shown, are firmly in the nuclear camp. But the anti-nuclear minority represents nearly a fifth of the members, who are effective in keeping up pressure on the Commission and member states.

Jasper Becker

East Germany banks on research

High technology will be the keynote of East Germany's economic strategy for the 1980s, First Secretary Erich Honecker told last month's Congress of the Socialist Unity Party. An intensive drive to modernize and automate industry, he said, should save the country some 2,800 million man-hours during the next five years, primarily through the application of microelectronics and an increase in the number of industrial robots to be introduced into industry during the period from an original estimate of 9,000 to some 40,000.

By 1985, the country should be virtually self-sufficient in the production of microprocessors and their components. Other priority areas will include fibre optics, laser technology, industrial applications of microbiology, hydrogen and biomass energy sources and the development of new sophisticated technologies to make the maximum use of raw materials at the minimum cost in energy. (The recent shortfall and future uncertainty of coal supplies from Poland have caused planners to pay renewed attention to the country's own lower-grade fossil fuels, mainly lignite).

Although Mr Honecker's review of the work of research establishments, from the Academy of Sciences downwards, stressed the technological and applied aspects, he did pledge the party to a systematic development of basic research "as the source of new discoveries on logical relations in nature and society". The univerisities, including their extension and evening courses for full-time workers, came in for special commendation as did the preliminary training courses for students at their future work-places.

The situation in the lower educational echelons, however, seems less satisfactory. Too few young people seem to be interested in scientific and technical subjects, and Honecker urged teachers and the media to make science and technology more attractive. **Vera Rich**

British countryside

Ever onward

The British government's Wildlife and Countryside Bill — regarded as the last opportunity for major legislation on nature conservation matters this century — is likely to generate as much controversy during its passage through the House of Commons as it did in the House of Lords earlier this year. The Lords considered a record 560 amendments to the bill and defeated the government on seven of them. Many of the arguments thrashed out then, on numerous topics relating to nature conservation, will be raised again, with the government trying to reverse some of the Lords amendments and various lobbyists trying to introduce more.

During the bill's second reading in the Commons last week, Mr Michael Heseltine, Secretary of State for the Environment, said that the government was still convinced that conservation would not be well served by compulsion. This is the chief bone of contention over clauses in the bill relating to the protection of sites of special scientific interest (SSSIs). In the Lords, the conservation lobby attempted — unsuccessfully — to amend clauses giving statutory protection to a few specially selected SSSIs to cover all such sites. The debate is likely to continue in the House of Commons committee appointed to debate the bill in detail.

The committee, which began sitting on 5 May, has 20 members, including Keith Hampson and Hector Munro for the

0028-0836/81/190005-02$01.00

government and Tam Dalyell and Stephen Ross for the opposition. As well as protection for SSSIs, other controversial topics include the banning of shooting on Sundays, allowing bulls on public footpaths and how to maintain marine nature reserves.

Judy Redfearn

Reagan's cuts

Handler objects

Washington

Dr Philip Handler, in his farewell address last week as president of the National Academy of Sciences, came close to accusing the National Science Board of cowardice in its failure to protest at the Reagan Administration's cut in the budget of the National Science Foundation (NSF).

The proposal is that the foundation's budget for 1982 be trimmed to $1,033 million, more than 30 per cent less than Mr Carter's request for $1,353 million. At its meeting in March, the board, which is formally responsible for the activities of NSF, approved a statement on the cuts in which it complained that it had been excluded from the budget decision.

The statement also mentioned "problems of serious concern" such as the need to upgrade university research equipment and its "limited capability to reprogram funds to take advantage of emerging research opportunities". There is, however, no direct criticism of the Reagan Administration's decision as such.

Dr Handler said plaintively last week "Would that that statement were not quite so subtle". Handler accepts that the board is constrained by its formal place within the Administration, but he told the academy that the board members' "statutory responsibility to establish the priorities and allocate resources had, in fact, been usurped". And he added: "In not saying so, in unequivocal language, they may allow this incident to become a precedent".

Handler admitted that for the members of the board to have been more forthright might have meant a confrontation with the Administration in which the National Science Foundation would be the loser. "But in crafting a statement in which protest is to be found only by sophisticated reading between the lines . . . the board seemed to be party to the actions of the Office of Management and Budget."

Handler also quoted the board's contention that it was difficult to be engaged in policy decisions about budget cuts in the light of the "economic emergency", which precluded the normal discussion with the board of NSF priorities. "Emergency? What emergency?" asked Handler. "The fact that NSF was given but 24 hours, when the board was not in town, to defend its right to support social science and science education, *inter alia*, is preposterous."

David Dickson

Nuclear power in Yugoslavia

Staying independent

Belgrade

Core loading is now under way at Yugoslavia's first nuclear power station in Krsko. Formally, the station — based on a Westinghouse light-water reactor — is years behind schedule, but this is in part due to an over-optimistic target of five years for construction.

The station has been built jointly by the republics of Slovenia and Croatia. The leading physics institutes of the two republics, the Rudjer Boskovic Institute (Zagreb) and the Josef Stefan Institute (Ljubljana), have been training staff for the new station while building was being completed, and working out the final safety report for the reactor.

Yugoslavia's commitment to nuclear power is strengthened by the presence of uranium deposits at Zirovski Vrh in Slovenia. There are few dissenting voices, notably hotel-keepers on the Dalmatian coast objecting to the proposed siting of the country's second reactor near the tourist resorts, and veteran anti-nuclear campaigner Dr Ivan Supek. Most Yugoslavs apparently see nuclear power as a way to greater energy self-sufficiency.

The TRIGA II reactor at the Josef Stefan Institute, used for heavy metal monitoring.

Before uranium mining and milling operations begin, however, a group from the Josef Stefan Institute is making a detailed survey of background radiation in the area. And this survey forms part of an international project to map the radio-ecology of the Danube basin.

Also being measured in the survey are heavy metal trace elements, including cadmium, arsenic and mercury. The area has a history of mercury mining and once the price of mercury picks up again the old mine is likely to be reopened. The mine closed in the mid-1970s and there is evidence of considerable mercury poisoning in the mine-workers involved. But studies of the effects of mining are being hampered by shortage of finance for a much-needed survey of the non-mining population in the mining area.

This highlights a general shortage of funds for research in Yugoslavia. There is no longer any central planning of science, and funds come either from direct contracts with industry, or else from a structure of "Self-Management Communities of Interest for Science" which administer funds voted from profits of individual commercial enterprises.

The Josef Stefan Institute gets 40 per cent of its funds from the Slovenian assembly of such communities and 60 per cent from contracts. But neither source has shown much interest in studying mercury levels in their region and monitoring a control population elsewhere. **Vera Rich**

Proton decay hunt

New Delhi

Indian and Japanese physicists have discovered three "candidate" proton decay events in a massive iron detector system set up 2,300 metres underground in the Kolar gold mines 100 km north of Bangalore in South India. Begun nearly five months ago, the experiment is to last two years during which at least six of these rare events are expected to be observed provided the proton lifetime does not exceed 10^{31} years.

The detector system consists of 140 tonnes of iron: 100 tonnes for the detector elements and the remainder for shielding. The detector elements are 4 m and 6 m long iron tubes of cross-section 10cm × 10cm. The 1,650 proportional counters are stacked to make a giant 35-layer cake that is 6m long, 4m wide and 4m tall.

The on-line electronics system and the detector configuration enable the total energy of the decay process, the decay vertex and a three-dimensional plot of the tracks of decay products to be obtained. There are some 10^{32} nucleons in the iron and assuming a detector efficiency of 50 per cent, it is hoped to reach a proton lifetime limit of 5×10^{31} years.

The main background to the decay events comes from cosmic rays. Neutrinos and some muons created in cosmic ray events in the atmosphere can penetrate down to the detector; but the Kolar equipment is so deep most of the muons are absorbed in the rock above. Nevertheless, by the end of last week Kolar had recorded 223 vertical muons; 8 horizontal muons, probably coming from neutrino interactions in the rock beside the detector; three neutrino events in the detector; and three events which so far can be explained in no other way than as proton decays. However, all three events have a track which reaches (or emanates from) the edge of the chamber, so they may yet be false events, mimicked by a particle entering the chamber from outside. If the events are confirmed, the corresponding proton lifetime would be $1–3 \times 10^{30}$ years, around the number being predicted by theories which unify the weak, electromagnetic and strong interactions into one grand "gauge theory with spontaneous symmetry breaking".

K.S. Jayaraman

0028-0836/81/190006 01$01.00

CORRESPONDENCE

Conservation sites

Sir — Since private landownership and private incentive are among the basic tenets of Western ideology, I have to agree with Muir (*Nature*, 12 March, p.82) that the populace must suitably entice landowners if it expects them to conserve rather than utilize their land. But it is a callous and immoral science that will not use funds to intervene in natural selection or in "the very evolution of the inanimate world" if these are observed to be proceeding in a deleterious direction. Surely the spending of funds on research directed towards new forms of medical treatment, or of flood or erosion control, is not an "affront to reason", bearing condemnation, and no-one is being misled by the professed importance of these or similar interventions.

Also, thankfully, I know of no science which claims that we can study anything in nature to the limits of its information content so that "nothing new will be learned by preservation". Such would be a conceited science indeed. And, a science that attaches no value to the existence of an object, save for its information content, shows a blatant contempt for existence itself. As Santayana has written (in *Reason in Science*), "If science deserves respect, it is not for being oracular but for being useful and delightful, as seeing is".

Unique usefulness and ability to delight are also properties of jewels or "gems" and of conservation sites, although Muir, in his derisive analogy, has conveniently overlooked these. He has also ignored the one striking difference between them: the virtual indestructibility of jewels as compared with the fragility of ecological systems.

So, must we look forward dispassionately to Muir's future world with its myriad rats, biting insects and noxious weeds, with its computer data banks full to overflowing, but without museums, zoos or conservation sites? Or, should we intervene to ensure a better future where usefulness and delight remain and where Muir's biologically deficient species are not extinct before their time?

A. H. Bretag

South Australian Institute of Technology,
Adelaide, South Australia

Literature search

Sir — PhDs normally take longer than three years to complete and their exact purpose is not agreed upon (*Nature* 22 January, p.217). As an SRC-supported research student, I would like to make a few relevant comments.

I have spent several months in my library carrying out the customary literature search. As time progresses, more and more literature has to be searched and one can only presume that the time to complete a PhD will increase! However, because I spent so long "doing" a search, I found cases of "published rediscoveries" which are attributable I think to: (1) a lack of time spent on a literature search, and (2) a language problem.

The time factor is overcome, at present, only by carrying out a minimal literature search (which in all likelihood will turn out to be inadequate) and by ignoring the language problem altogether by ignoring the literature published in a foreign language. The dilemma facing a research student is obvious. If he/she tries to carry out a thorough literature search and learn a new language then the time spent doing this bites into those three years of support. At a recent course which I attended, which was specifically designed to teach scientists how to translate scientific Russian or German, only five people attended. The conclusion is (from albeit scanty evidence) that researchers choose to ignore the language problem.

Regarding actual literature searching it is only recently that I have found out about "computerized" abstracts, although, I'm told, they have been available for a few years. I would urge the SRC to look into the use (or lack of use) of "computerized" abstracts, if they are not doing so already, as such facilities will become increasingly important and have the potential for saving a lot of time.

I suggest that it would be beneficial if research students were given, from the outset, a course on library and information science so that they are aware of all possible sources of information. I would also suggest that it be made compulsory for research students to learn a foreign language relevant to the scientific discipline. These proposals may take up money and indeed time but to me, at least, appear a necessity if good efficient research is to be carried out. On reflection, what I have suggested favours a more "training for research" type of PhD.

Andrew Brooks

Department of Astronomy,
University of Edinburgh, UK

Avoiding fraud

Sir — The recent alarms about fraud in American science raise the more general problem of quality-control in research. As you correctly observe (*Nature* 9 April, p.433), peer review must be supplemented by "full and frank discussion within individual laboratories". But the process cannot stop there. When I analysed this problem in my book (*Scientific Knowledge and its Social Problems*, Oxford University Press, 1971), I remarked on the inherent difficulty of the quality-control operation in science, since there is no possibility of creating an external inspectorate for assessing the products of the complex and subtle craft skills of science.

Research is therefore necessarily a largely self-regulating activity in this respect, accomplished by peer-review and journal refereeing; but the problems of iteration, "who guards the guardians?" are then even more severe. My conclusion was that integrity and morale, at the highest professional levels, are more crucial to the health of science than perhaps in any other organized social activity.

The various moral imperatives of science, propounded by Merton, Popper and Polanyi, can thus be seen to be very relevant to the survival of worthwhile science. However, they are not so much *a priori* definitions of "science", as descriptions of attitudes and commitments whose presence and effectiveness are entirely contingent. The variability of standards of quality between fields and between milieux bears this out.

When so much of scientific information relevant to public policy is now produced not as "public knowledge" but rather as "corporate knowhow" (in state bureaucracies or private firms), the traditional ideals and norms of science lack an appropriate social context for their reinforcement and maintenance. Under these circumstances it becomes implausible to maintain that only a prejudiced or malcontented opposition can doubt the factual veracity of any piece of technical information used as official testimony in policy debates.

J.R. Ravetz

Department of Philosophy,
University of Leeds, UK

Cladistic clues

Sir — It is unfortunate that the attempts of Dr Miles and the Department of Public Services of the British Museum to display the most logical basic level of analysis of organisms (the level of relative organism-to-organism comparison) is misinterpreted as an attempt to coerce the general public into accepting "a fundamentally Marxist view of life". As one may perceive from the series of replies to L. B. Halstead and the cladistic literature, cladistics is a method which links many very different individual biologists working from very different sets of preconceptions. However, they seem at least to share the opinion that we learn something about organisms by studying what they are made of. The cladist may converge on the Marxist idea that events occur by "taking the form of a leap from one state to another." The cladist may similarly converge with the catastrophist creationist. However, the cladist's stand can be just the opposite of a dogmatic stand, that is, it can be an admission that there are not enough biologists to fill in all the holes in our knowledge of the history of life, and that our inadequate sample of the history of life allows us to see so little "gradualism" that one cannot dogmatically claim that it is the rule.

Halstead's point that public scientific institutions are accountable to the public is well taken. However, publicly employed scientists are accountable to the public for their salaries with ideas and interpretations which the public at large does not have time to make, as well as being accountable for actual specimen displays. Like any curator in history responsible for exhibitions, Dr Miles shares some ideas with scientists of his day, and has some ideas of his own, all of which are reflected in the exhibits he edits and presents to the public. Like any editor, he is forced to make editorial decisions; as one who studies the taxonomic group that Dr Miles has spent most of his career studying, I see no more competent individual than himself.

However, I share some of S. J. Gould's romanticism for great old halls which are closed and renovated into something that is a total stranger to me. Museums cannot help but be places of accumulation, and this catastrophistic kind of replacement of the old with the new does not seem to be the most accurate portrayal of the museum's progress in

Increased majority

In the correspondence "Majority verdict" published in *Nature* 30 April, p.730, R.L. Batten should be added to the list of signatories.

research. Perhaps the most instructive exhibition policy would be to preserve some small corner of every old exhibition hall which would allow the public to see that which their predecessors saw, and which would allow the museum to see what it stood for in the public eye. In the next generation of British Museum exhibits, maybe we will see evolution not as untestable, unscientific theory, but as a more complex though rather less testable theory than cladistic relationship, and as complex and considerably more testable than creationism.

W. Kohlberger

Department of Geology and Geophysics,
Yale University, New Haven,
Connecticut, USA

Cafeteria feeding

Sir — A. S. Cole (*Nature* 2 April, p.356) objects to the phrase "cafeteria feeding", applied to experimental selection of "palatable items of food" by rats. I disagree: it seems an unobjectionable, indeed mildly amusing, and appropriate piece of jargon — for once, easily understood. The only criticism might be that, in the original usage, the *palatability* of items on offer is not a *sine qua non.*

J. R. Baker

NERC Culture Centre of
Algae and Protozoa,
Cambridge, UK

Ten years of natural history

Sir — In 1972, the trustees of the British Museum (Natural History) approved a proposal prepared by a panel of the museum's scientific staff, chaired by the then director, for a new approach to the visiting public. This proposal sought to replace the then existing exhibition which rigidly reflected the divisions of the museum into the five separate scientific departments of Botany, Entomology, Mineralogy, Palaeontology and Zoology. The new approach was timely because it coincided with, *inter alia,* changes in biology teaching in schools and an increasingly sophisticated treatment given to natural history on television.

The aims of the new exhibition scheme were summarized as follows: "To present the public with an integrated view of modern biology. All known forms of life will be included, but the new exhibition will do much more than merely demonstrate the diversity of nature. It will show how living things interact with each other and with their physical environment; it will describe the chemical and physical processes that keep them alive; it will introduce the concepts of heredity and evolution, of ecosystem and energy cycles; and it will examine man's role in the living world[1]." A long-term aim was, and still is, to include a greater number of organisms from a wider range of groups than had been shown in the past, combining both recent and fossil organisms in a single evolutionary arrangement.

The trustees have maintained a continuous and close interest in the implementation of this policy, and have conducted a careful review from which they concluded that the museum's Public Services Department has made a bold and largely successful start on implementing the long-term plan. The trustees now consider it appropriate to outline their plans for the next ten years so that those who visit and use the museum may be aware of changes and of the underlying rationale. The trustees have concluded that the scheme should proceed as shown in the table below.

The trustees believe that the implementation of this programme will be beneficial. It will, for example, allow fossil and recent mammals to be brought together in a single rational system for the first time in the museum, and it will not only give greater space to the insects but it will also relate them to the other arthropod groups.

In addition, the Department of Public Services is developing a proposal for an exhibition devoted to British Natural History to be brought forward to the earliest feasible time. This gallery would be designed to meet the needs of committed naturalists to whom the trustees acknowledge a responsiblity, although the prime commitment of the exhibitions must be to the visiting public in general.

The new stages of the exhibition proposed above will have much in common with the older galleries because they will concentrate on the museum's traditional subject matter, the diversity of living organisms. Eventually diversity exhibits are likely to comprise about two-thirds of the completed exhibition scheme. Some parts of the existing displays from, for example, British Insects and Fossil Mammals, will be used in the new exhibitions.

The trustees are also aware of the need for an exhibition related to the museum's interests in geology and have accepted in principle proposals by a Joint British Museum (Natural History)/Natural Environment Research Council Study Group for a co-ordinated display of minerals, gemstones, rocks and meteorites in the BM(NH) and the Geological Museum.

These proposals will necessarily involve a considerable upheaval in the public galleries during the next decade. In order to minimize the effects of this disruption: (1) A temporary display will be provided when any major group of organisms is not represented in a permanent display for a substantial period of time. (4) Access to the Whale Hall will be maintained even though work may be in progress there; thus the Blue Whale will be continuously on display to the public. (3) The larger fossil mammal specimens will be inaccessible only during the period of transit from one gallery to another, as was the case with dinosaurs. And (4), the insects will be off display for a matter of weeks only.

Changes of this magnitude will cause many difficulties and the museum welcomes constructive suggestions, either now or as the work proceeds. The trustees are aware of the controversy aroused by the first phases of the new policy. However, a paper by Dr G. C. S. Clarke and Dr R. S. Miles[2] of the museum's Department of Public Services explained the ideas behind the changes which had provoked the controversy and invited constructive criticism. No replies were received. The museum has continuously monitored the effectiveness of the new exhibition and the responses of the visiting public and has good evidence that the vast majority of the visiting public, especially schoolchildren and their teachers, greatly appreciate the new exhibitions[3]. Suitable modifications have been made when they would lead to improvements.

To assist visitors to understand the complicated changes which are taking place the museum is preparing a small temporary exhibit of the plans.

T. R. E. Southwood
(*Chairman of the Board of Trustees*)
R. H. Hedley
(*Director*)

British Museum (Natural History),
London SW7, UK

Exhibitions planned in the next ten years

Subject	**Location**
Origin of Species (to be opened in May 1981)	Upper South-West Gallery
Mammal diversity (1) — cetaceans, artiodactyls, perissodactyls, hyraxes, sirenians, elephants, rodents and related fossil orders	Whale Hall
Mammal diversity (2) — mammal-like reptiles, monotremes, marsupials, edentates, carnivores, insectivores, bats and fossil groups	Galleries now used for insects and marine invertebrates
Mammal diversity (3) — primates	West balcony of Central Hall
Arthropod diversity (1) — insects centipedes, millipedes and Onychophora	East Pavilion and South-East Gallery
Arthropod diversity (2) — Crustaceans	South-East Gallery
Arthropod diversity (3) — spiders, scorpions, ticks, mites and trilobites	South-East Gallery
Unity in diversity — an introduction to all the museum's exhibitions	Entrance to South-East Gallery

1. *Report of the British Museum (Natural History), 1972–74*, p.3 (Trustees of the British Museum (Natural History), 1975).
2. Clarke, G. C. S. & Miles, R. S. *Biologist* **27**, 81–83 (1980).
3. Alt, M. B. *Mus. J.* 80, 10–19 (1980); also unpublished reports on the school use of the Natural History Museum, 1979 and 1980.

0028-0836/81/190008-01$01.00

Spatial frequency analysis in vision

from Oliver Braddick

READERS of *Nature* will probably be aware that many experimenters and theorists in visual perception and physiology have taken to using the language of Fourier analysis: they test vision with sine wave grating patterns and discuss the behaviour of the visual system in terms of spatial frequency response. Why has this approach become popular, and has it been justified as enhancing our understanding of vision?

Descriptions in terms of frequency are most familiar when speaking of periodic events in time, such as sound or radio waves, but they may be used for spatial patterns in an analogous way. The number of repetitions of a pattern within a unit of distance is its spatial frequency. The mathematics of Fourier analysis which allow any complex sound to be described as a set of harmonics also allow any visual pattern to be described as a sum of spatial frequency components, that is, patterns whose luminance varies across space as a sine wave. A system designed to transmit time-varying signals, such as a hi-fi set, commonly has its performance specified by its frequency response. Similarly, how well a system transmits spatial patterns can be characterised by its response to different spatial frequencies. In optics, this has been a commonplace technique for evaluating lens systems for many years. The performance of the human eye, however, is determined not only by its optical quality in forming the retinal image, but also by how well the visual system signals the information in that image. It therefore seemed natural to extend the idea of measuring spatial frequency response to the whole system[1,2]. This can be done if we make the fairly uncontentious assumption that for a human subject trying to detect a low-contrast grating pattern, a low threshold indicates a high response of the visual system to that spatial frequency, and so on, using thresholds to measure the contrast sensitivity function.

The spatial frequency response of some component within the visual system may also be measurable: this can be done directly in physiological experiments by measuring the response of single nerve cells to gratings of various frequencies, and indirectly in psychophysical experiments by testing whether two gratings interact in their effects on perception. For example, if one grating can mask detection of another, it is concluded that they must be processed by a common neural 'channel'[3].

The flood of experimental work in what has been dubbed 'gratingology' has come largely from the idea proposed by Campbell and Robson[4] that Fourier methods were valuable, not just for characterising the performance of the whole visual system, but also for understanding the way the system is divided into parallel information-carrying channels. It is generally believed that the properties of these channels define the 'language' in which visual information is conveyed to, and perhaps within, the brain. The idea that this language is best analysed in terms of spatial frequency has been vigorously asserted and contested. If we are to avoid this controversy obscuring the significance of the experimental results, we must keep clear the distinctions between various levels at which spatial frequency methods and Fourier theory may be applied to vision.

First, the spatial frequency response of the visual system can be used as a valuable descriptive measure, without commitment to any particular assumptions about the properties of the visual system. Second, insofar as the visual system, or some component of it, can be treated as linear, the spatial frequency response has important predictive power: since any pattern can be considered as the sum of sine wave components, the visual response to any pattern can be predicted by summing the responses to the separate components. Third, it has been proposed that the visual system at some level consists of parallel, independent channels, each responding to its own limited band of spatial frequency. If so, the visual system itself is performing, at least crudely, an analysis of the visual pattern into its frequency components. Fourth, and most radically, it is possible that this division of the visual pathway into spatial frequency channels might indicate a fundamental way in which visual information is organized: that is, perceptual processes such as selective attention, or recognition, might operate on representations of the visual world in terms of spatial frequency components rather than the more familiar kind of representation in terms of spatial location.

The descriptive use of spatial frequency analysis, at its simplest, amounts to saying that it is important to know not only how well vision conveys very fine detail (high spatial frequencies), but also how well it conveys information about the patterning of light and dark in the image at every scale from the very coarse to the very fine. One application of this is clinical: the quality of vision is conventionally measured in terms of acuity (that is, detail resolution) but some patients complain of poor vision even when their measured acuity is normal. In many cases they can be shown to have abnormally low sensitivity to spatial frequencies below the acuity limit. Indeed, differences in the pattern of contrast sensitivity may be diagnostically significant[5]. An important side effect of the descriptive use of spatial frequency response has been to emphasise the usefulness of contrast as a parameter of vision. Even for non-sine wave stimuli, contrast threshold is often a more useful measure than luminance threshold, since contrast can be varied without introducing the radical changes that occur in the visual system with adaptation to different light levels.

If spatial frequency response is found to have predictive as well as descriptive power, that implies that the visual process concerned shows linear spatial summation. This in turn implies that it could be described equally completely in spatial as in spatial frequency terms. This is true, for instance, of one class of retinal ganglion cells (X-cells)[6]: their response to a pattern can be predicted either from the summation of response to its spatial frequency components, or from summation of the local responses given by each point in the receptive field. The spatial sensitivity profile of the receptive field and the spatial frequency response function may be

---

Oliver Braddick is in the Department of Experimental Psychology, University of Cambridge.

0028-0836/81/190009-03$01.00

mathematically transformed into each other. In experimental practice, this complete description may be more readily achievable by testing the spatial frequency response, since the cell's response to a grating which covers the whole receptive field is stronger and clearer than that to illumination of a small spot at some point within it.

It can be argued, though, that the spatial frequency approach has a particular value when the requirement of linearity is *not* met. This is because frequency analysis provides tests that very clearly indicate the presence and nature of any non-linearity present. Thus, it was through the investigation of response to sinusoidal stimuli that an important class of non-linear (Y) retinal ganglion cells was discovered by Enroth-Cugell and Robson[6], and that the nature of the mechanisms which produce their characteristic behaviour has begun to be understood[7].

Taking the performance of the visual system as a whole, there are clearly various significant non-linearities, but for the conditions under which many detection experiments are done linear predictions work quite well[4]. This means that it is useful to consider the separate visual response to each of the frequency components making up a complex pattern. Using this mathematical analysis does not require us to suppose that the visual system breaks up the pattern into these components in any physical sense. The frequency response can be taken to describe the properties of visual system as a homogeneous linear spatial filter — the 'single channel' model. The alternative is a 'multi-channel' model in which the analysis into spatial frequency components has more physical reality. In this model, the overall frequency response is the envelope of the sensitivity functions of many independent channels, and different channels may respond to components in a complex pattern. A wide range of psychophysical experiments support the idea of multiple spatial-frequency channels[3,8], although there are still considerable quantitative problems to be resolved on how many parallel channels there are[9], and how narrow is the band of spatial frequencies to which each responds[10].

Physiological experiments, too, show that the visual cortex contains cells responding to different bands of spatial frequencies in the same region of the visual field[11]. It is true that, insofar as these cells show linear spatial summation (as many of them do), their receptive field sensitivity profiles equally describe their properties[12]. For fields of similar form, the differences in optimal frequency will correspond to differences in receptive field size. Even so, the frequency description has a good deal of heuristic value. For instance, a qualitative account in terms of receptive field dimension might lead one to expect that such cells would differ strikingly in the width of bright (or dark) bar to which they respond best. In fact, this turns out to characterise their differences much less clearly than does their spatial frequency response[13].

However, spatial frequency response alone does not give all the information expressed in a spatial sensitivity profile. This is because the response is a function of the phase as well as the spatial frequency of a grating. A cell which responds well to bar patterns will show optimal response to different grating frequencies when they have a peak of the waveform in a given position, while responsiveness to edge patterns corresponds to the requirement that waveforms of different frequency should have their maximum gradient in the same position. What is apparent from making the distinction in terms of phase, rather than edges versus bars, is that any pattern can be considered a mixture of the two phase relationships. Both psychophysicists and physiologists[12] are now exploring the possibility that the division into two types of phase response is another form of 'channeling' for encoding spatial patterns in vision.

Given the existence of channels that respond independently to different spatial frequency bands, should we conclude that visual perception operates in the Fourier domain? In modern digital image processing, it is often found useful to perform various operations on an array of, say 256×256 spatial frequencies, each of which is derived in the Fourier transform of an entire picture of 256×256 local elements. It is most unlikely that visual perception works in anything like this way. Physiological and psychophysical evidence concur that, at the levels we have been considering, analysis is fairly local: receptive fields are limited in size. There is a mathematical inverse relationship between the width of the spatial frequency band to which a channel responds and the range of spatial positions to which it responds[14]. It appears that the visual system represents a compromise, dividing the spatial frequency spectrum fairly coarsely within local, but not very small, areas.

In fact, it would probably be pointless for the visual system to measure spatial frequency components, or anything else, over the visual field as a whole. What we usually want to recognise is not the visual field as a unified entity, but objects within it. Could we be recognising an object by means of the spatial frequency components detected by a local set of channels? In most cases, this seems quantitatively unlikely. Analyses of spatial frequency channels in central vision have generally found them only for spatial frequencies above 1 cycle/degree. Such channels could encode an adequate frequency spectrum only for an object that was roughly ½ degree across, or smaller. With the important exception of written characters, we need to recognise most objects over a range of visual angles markedly larger than this. A similar conclusion arises from physiological receptive field sizes, and from the extent of visual field analysed by a single cortical 'hypercolumn'[15]: at the level of visual cortex, a single object will normally be analysed by many, separate, receptive fields. An important aspect of the processing of visual shape must be the 'grouping' of information from local analysers spread out over an area of the visual field.

What, then, might be the role of spatial frequency analysis in the process of perception? Marr and Hildreth[16] have suggested that coincident signals in different frequency channels provide an effective criterion for finding edges in the visual field. In this view spatial frequency analysis is used at a rather low level to define features, which themselves are part of a representation in the domain of space rather than spatial frequency. Separate processing of spatial frequency bands has been proposed as playing a rather similar role of local disambiguation in the computation of stereo disparity[17,18]. In these models, the information signalled by spatial frequency channels is not separately available for higher perceptual processes. This argument has been supported by experimental evidence that patterns can differ quite markedly in the Fourier domain without it being possible to use these differences for perceptual grouping[19] (although this result concerns the ability to separate componets that differ in orientation, not in spatial frequency itself).

On the other hand, one of the best known effects in grating psychophysics is the 'perceived spatial frequency shift' produced by adaptation to a grating[20]. This

---

1. Schade, O.H. *J. opt. Soc. Am.* **46**, 721 (1956).
2. Campbell, F.W. & Green, D.G. *J. Physiol., Lond.* **181**, 576 (1965).
3. Braddick, O., Campbell, F.W. & Atkinson, J. in *Handbook of Sensory Physiology, Vol. VIII: Perception* (eds Held, R., Leibowitz, H. & Teuber, H.L.) (Springer, Heidelberg, 1978).
4. Campbell, F.W. & Robson, J.G. *J. Physiol., London.* **197**, 551 (1968).
5. Arden, G.B. *Br. J. Ophthalmol.* **62**, 198 (1978). Bodis-Wollner, I. & Diamond, S.P. *Brain* **99**, 695 (1976). Hess, R.F. & Howell, E.R. *Vision Res.* **17**, 1049 (1977).
6. Enroth-Cugell, C. & Robson, J.G. *J. Physiol., Lond.* **187**, 517 (1966).
7. Hochstein, S. & Shapley, R.M. *J. Physiol., Lond.* **262**, 237 (1976).
8. Blakemore, C. & Campbell, F.W. *J. Physiol., Lond.* **203**, 237 (1969). Graham, N. & Nachmias, J. *Vision Res.* **11**, 251 (1971). Graham, N., Robson, J.G. & Nachmias, J. *Vision Res.* **18**, 815 (1978). Sachs, M., Nachmias, J. & Robson, J.G. *J. opt. Soc. Am.* **61**, 1176 (1971).
9. Wilson, H.R. & Bergen, J.R. *Vision Res.* **19**, 19 (1979). Watson, A.B. *Vision Res.* (in the press).
10. Graham, N. *Vision Res.*, **17**, 637 (1977). King-Smith, P.E. & Kulikowski, J.J. *Vision Res.* **21**, 235 (1981).
11. Maffie, L. & Fiorentini, A. *Vision Res.* **13** 1225 (1973).
12. Movshon, J.A., Thompson, I.D. & Tolhurst, D.J. *J. Physiol., Lond.* **283**, 53 (1978).
13. De Valois, R.L., Albrecht, D.G. & Thorell, L.G. in *Frontiers in Visual Science* (eds Cool, S.J. & Smith, E.L. III) (Springer, New York, 1977).
14. Mackay, D.M. *Nature* **289, 117 (1981).**
15. Hubel, D.H. & Wiesel, T.N. *Proc. R. Soc. Lond.* **B198** (1977).
16. Marr, D. & Hildreth, E. *Proc R. Soc. Lond.* **B207**, 187 (1980).
17. Marr, D. & Poggio, T. *Proc. R. Soc. Lond.* **B204**, 301 (1979).
18. Mayhew, J.W.W. & Frisby, J. *Perception* **9**, 60 (1980).
19. Mayhew, J.E.W. & Frisby, J.P. *Nature* **275**, 438 (1978).
20. Blakemore, C., Sutton, P. & Nachmias, J. *J. Physiol., Lond.* **210**, 727 (1970).
21. Campbell, F.W. & Marr, D. Discussion in *Phil. Tans. R. Soc.* **B209**, 9 (1980).

0028-0836/81/190010-02$01.00

effect is best explained by the assumption that perceived spacing of the grating is directly determined by which spatial frequency channel responds most strongly. It can be argued that judging the spatial frequency of a grating is more related to recognising a repetitive texture, in which we are unconcerned with the spatial identity of individual features, than to perceiving the size and shape of a discrete object. Maybe objects are represented in spatial terms, but the textures of their surfaces in spatial frequency terms.

Speaking of 'objects' should not lead us into the assumption that there we divide any particular visual scene in a fixed way. Clearly for some purposes we are primarily concerned with the gross layout of our visual environment, while for others we need information that is present in fine detail. It is tempting to suppose that for different perceptual tasks we might switch our attention to different spatial frequency bands. However, it is hard to find an experimental example that demands an explanation as spatial frequency tuning of attention, and there are certainly instances where we do *not* find it possible to select a frequency band in this way[21].

Many visual scientists who are sceptical or open minded on how far higher perceptual processes speak the language of spatial frequency nonetheless find that language invaluable for describing and manipulating their experimental stimuli and the information signalled in the visual system. New insights come from thinking of the visual image as containing patterns of light and dark over a range of spatial scales, or 'degrees of blurredness'. Formally, this idea, and analyses that embody it, may be expressible as exactly in spatial as in spatial frequency terms. The fact remains that by exploring the implications of the spatial frequency description, and by using stimulus patterns which were simple to describe that way, we have learned a great deal about vision and by all indications have more to learn yet. □

Sliding charge density waves

from P. A. Lee

THE idea of a charge density wave transition goes back to Peierls[1] and, independently, to Fröhlich[2] in 1954. They reasoned that a one-dimensional metal can lower its energy by distorting the lattice and forming a gap at the Fermi surface, thereby making a transition into an insulating state. The term 'charge density wave' refers to the fact that the lattice and the electron charge density form a new periodic structure, with a wavelength λ that is longer than the original lattice period a. Fröhlich further reasoned that if the charge density wave is incommensurate, that is, a/λ is not a simple rational fraction like 1/2 or 2/3, then the entire charge density wave structure can slide through the lattice, thereby contributing to the electrical conductivity. Since these ideas were originally formulated in terms of one-dimensional systems, for many years they were considered with nothing more than theoretical curiosity.

In the early 1970s the interests of many physicists turned to quasi one- and two-dimensional systems. The existence of a charge density wave ground state in these systems turns out to be the rule rather than the exception, and a large number of examples have been discovered. Experimentally the onset of the charge density wave is signaled by a rise in the resistivity at some temperature T_c. In quasi one-dimensional systems a transition to an insulating state often occurs, whereas in the layered compounds the system often stays metallic. The most direct microscopic evidence for a charge density wave transition is the appearance of new superlattice spots associated with the new lattice periodicity λ which is detected by X-ray, neutron or electron diffraction. Prominent examples of charge density wave systems are the organic charge transfer salts[3], such as tetrathiafulvalene-tetracyanoquinodimethane (TTF-TCNQ) and layered compounds[4] formed from transition metals and chalcogens, such as TaS_2 and $NbSe_2$.

However, until 1976, none of the charge density wave systems had exhibited the sliding conductivity envisioned by Fröhlich. This is because the charge density wave was either commensurate or pinned by impurities. Impurities are sensitive to the charge density oscillations and can lock the charge density wave in place, a phenomenon very analogous to static friction[5]. As in friction, a sufficiently large force should dislodge the charge density wave. Thus the observation[6,7] in 1976 of nonlinear electrical conductivity in $NbSe_3$ was greeted with a good deal of interest. In $NbSe_3$ there are two transitions at 59K and 144K; below each of these temperatures the resistivity rises. It was found that beyond a certain threshold electric field[8], which can be as low as a few meV per cm for the low-temperature transition and 0.1 eV per cm for the higher-temperature transition, the resistivity becomes nonlinear and begins to drop, until it saturates to a value close to that expected if the charge density wave transition had not occurred.

The conductivity is also found to be highly frequency dependent, rising with increasing frequency until it coincides with the high field d.c. value at ~ 100 MHz (refs 7 and 9). It soon became clear that the electric field is not destroying the charge density wave, because an X-ray diffraction experiment shows that the superlattice spots are not affected by a current through the sample even when the current exceeds the linear threshold[10]. Diffraction experiments reveal incommensurate lattice distortions with periods λ such that $a/\lambda = 0.24$ and 0.26 for the upper and lower transitions. These data are consistent with an interpretation in terms of impurity pinning of the charge density wave[11]. The charge density wave can undergo small oscillation, thus accounting for the frequency-dependent conductivity. The pinning frequency is very low, possibly because the impurity concentration in $NbSe_3$ is very small. What makes $NbSe_3$ special among charge density wave systems in this respect is not entirely understood at present. The weakness of the pinning also manifests itself in the low-frequency dielectric constant, which has been measured to be greater than 10^8 (see ref. 9). Presumably a small electric field is sufficient to overcome the weak pinning so that the charge density wave can slide in the way envisioned by Fröhlich — an interpretation supported by an experiment in which impurities were introduced into $NbSe_3$ and the threshold electric field was found to increase[12]. Instead of classical depinning, an alternative mechanism involving the quantum mechanical tunneling of macroscopic segments (up to several microns in length) of charge density waves has been proposed[13], but all existing theories are based on the collective motion of the charge density wave ground state described by Fröhlich.

Further study of $NbSe_3$ yielded more surprises. When the current exceeds the threshold, voltage noise across the sample greatly increases. Superimposed on this background noise are periodic components with well defined frequencies[8], which scale linearly with the excess nonlinear current[14] and are typically in the range of kHz to MHz. A model for the noise is that it arises when the charge density wave moves through the impurities or the lattice, much like a particle rolling down a washboard. If certain assumptions are made about the charge being carried by the sliding charge

---

P.A. Lee is at the Bell Laboratories, Murray Hill, New Jersey.

1. Peierls, R. E. *Quantum Theory of Solids* (Oxford University Press, London, 1955).
2. Fröhlich, H. *Proc. R. Soc.* **A223**, 296 (1954).
3. *The Physics and Chemistry of Low Dimensional Solid* (ed. Alcacer, L., 1979).
4. Wilson, J. A. *et al. Adv. Phys.* **24**, 117 (1975).
5. Lee, P. A. *et al. Solid State Commun.* **14**, 703 (1974).
6. Monceau, P. *et al. Phys. Rev. Lett.* **37**, 602 (1976).
7. Ong, N. P. & Monceau, P. *Phys. Rev.* **B16**, 3443 (1977).
8. Flemming, R. M. & Grimes, C. C. *Phys. Rev. Lett.* **42**, 1423 (1979).
9. Grüner, G. *et al. Phys. Rev. Lett.* **45**, 935 (1980).
10. Fleming, R. M. *et al. Phys. Rev.* **B18**, 5560 (1978).
11. Lee, P. A. & Rice, T. M. *Phys. Rev.* **B19**, 3970 (1979).
12. Ong, N. P. *et al. Phys. Rev. Lett.* **42**, 811 (1979).
13. Bardeen, J. *Phys. Rev. Lett.* **45**, 1978 (1980).
14. Monceau, P. *et al. Phys. Rev. Lett.* **45**, 43 (1980).
15. Fung, K. K. & Steeds, J. W. *Phys. Rev. Lett.* **45**, 1696 (1980).

0028-0836/81/190011-02$01.00

density wave (a quantity under debate at present), the frequency of the noise can be accounted for[14].

Recently there was a report of dark-field electron microscopy on NbSe$_3$ (ref. 15). An image of an area of the sample is formed using the superlattice reflection alone; thus any feature in the imaging is due to the charge density wave formation. It is found that stripes approximately 200Å wide run along the conducting axis, and across the stripes are bands of lighter and darker intensities, suggestive of a Moiré pattern. It appears that the charge density wave structure may itself be broken into domains and that more than one superlattice wave vector is possible, so that interference between them forms the Moiré pattern. What is even more intriguing is that at low temperatures the bands are found to shift in time, so that the entire pattern shimmers on a time scale of a fraction of a second. This suggests some time-dependent motion of the charge density wave, but it is occurring on such a slow time scale that its connection with the nonlinear and frequency-dependent conductivity is unclear.

It is rare to find a material like NbSe$_3$ that exhibits such a rich variety of unusual properties. At present the idea of a sliding charge density wave provides a framework for the qualitative understanding of many of these properties, but clearly much work remains to be done before a quantitative understanding is achieved. □

Mobile genetic elements: new Soviet studies

from A. D. Mirzabekov

A CONFERENCE devoted to studies of eukaryotic mobile dispersed genetic elements was held on November 14, 1980 as a session of the seminar 'Chromosoma' at the Institute of Molecular Biology (USSR Academy of Sciences, Moscow).

In the *Drosophila melanogaster* genome mobile dispersed genetic (mdg) elements are multiple, scattered throughout the whole genome, and can move around. They are also actively transcribed (G.P. Georgiev*). At present, about 20 families of mdg elements have been described and comprise as much as five per cent of the whole *D. melanogaster* genome. The variable location of mdg elements is the most striking example known of genomic differences between individuals of the same species.

Y.V. Ilyin[1] described the structural organization of the two elements, mdg1 (a ~ 7kilobase (kb) element including the previously described Dm225 and Dm234 sequences) and mdg3 (~ 5 kb) cloned in the plasmid pBR322. In cell culture, mdg1 and mdg3 are amplified but not tandemly repeated. Like all other transposons, mdg1 and mdg3 contain long terminal repeats (LTRs) at the ends of the gene.

Sequence data show the structure of the LTRs of mdg3 (Bayev *et al. Nucleic Acids Res.* **8**; 3263, 1980; A.S Krayer, V. Kulguskin*) to be perfect direct repeats 268 bp long flanked by two mismatched 18 bp inverted repeats. The complete mdg 3 element is then bounded by short (5 bp) direct repeats probably derived from host target sequence duplication. Mdg1 contains longer perfect direct repeats (442 bp) flanked by mismatched 16 bp inverted repeats. In this respect, the organization of mdg1 LTR is similar to LTRs of the *copia* gene (Levis *et al. Cell* **21**; 580, 1980) or terminal repeats of integrated proretroviruses (Beveren *et al. Proc. natn. Acad. Sci. U.S.A.* **77**; 3307, 1980; Dhar *et al.* ibid; 3937, 1980). Within both LTRs, signal sequences for transcription initiation (Hogness box) and polyadenylation (AATAAA) are found.

The whole mdg sequences of mdg1 and mdg3 are transcribed, initiation and termination sites being localized within LTRs (V.G. Chmeliauskaite*). Both DNA strands are transcribed giving rise to the long double-stranded RNA stretches found in total poly(A)$^+$ RNA from cell culture. However, the efficiency of transcription of one strand is much higher than that of the other. At least in *D. melanogaster*, hybridization to very long double-stranded RNA can be used as a diagnostic test in the search for new mdg elements. Besides full-length transcripts, spliced poly(A)$^+$ RNAs were found among mdg1- and mdg3-specific RNA sequences. The major part of mdg transcripts are located in free ribonucleoprotein particles in the cytoplasm and only a minor fraction is bound to polysomes.

M.D. Golubovsky (Institute of Cytology and Genetics, Novosibirsk) showed that the frequency of an unstable mutation in the *sn* locus of *D. melanogaster* can simultaneously increase in geographically distant natural populations of flies. This may be due to 'infection' with a movable element. In addition, a mutation was found which influences both *sn* and *ct* loci and probably depends on the presence of the same element. A correlation between instability in *white* locus and mdg 1 insertion was shown by E.S. Belyaeva†.

The localization of the mdg elements in chromosomes appears highly conserved during differentiation as they have been found in the same position in cells of salivary glands and stomach (E.V. Ananiev†). They are also stable in inbred lines of *D. melanogaster*. The probability of 'jumping' for mdg1 and mdg3 was estimated to be of the order of 10^{-5}.

Drastic changes in mdg1 and mdg3 localization were found to take place during back selection for high sexual activity among inbred flies characterized by low sexual activity (V.S. Gvozdev†). Such transpositions might facilitate adaptation of the organism to the environment through influence on the expression of other genes.

Mdg elements appear species specific: those of *D. melanogaster* cannot be detected in *D. virilis* and vice versa (M.B. Evgen'ev*). A new mdg family has been cloned by G. N. Yenikolopov from *D. virilis* genome. *D. littoralis* chromosomes included in the genome of *D. virilis* are able to accept *D. virilis* mdg copies without crossing-over.

A.N. Gudkov (Oncological Centre, Moscow) discussed endogenous proretroviruses and their similarity to mdg elements. Different strains of hen differ strikingly in the number and localization of proviral sequences as studied by Southern blotting analysis. The existence of a natural strain completely lacking RAVo-like sequences was described. Together with the data of Astrin *et al.* (*Cold Spring Harb. Symp. quant. Biol* **344**; 1105, 1979) this suggests that the presence of proviral sequences is not essential for the organism.

Sequences homologous to the oncogene of mouse Moloney sarcoma virus (putative protooncogene) and to some other sequences of the virus (human C-type proretrovirus) are found in the normal human genome (E.A. Zabarovsky*). The protooncogene-containing sequence has been cloned from the human genome.

D.M. Kramerov* described the properties of two mdg-like elements cloned from the mouse genome selected on the basis of efficient hybridization to long stretches of double-stranded RNA (dsRNA-A). One of these cloned sequences was found to be responsible for the synthesis of intracisternal A-particles. Also, endogeneous C-type proviruses were detected among dsRNA-A-hybridizing sequences of mouse genome.

Structural similarities between mdg elements of *Drosophila* and endogenous proretroviruses were stressed by Georgiev. He also discussed a 'promoter' hypothesis of carcinogenesis (see *J. theor. Biol.* **25**;473, 1969) which considers oncogenic transformation to result from a potential cellular oncogene falling under the control of the viral promoter. Mdg elements may provide an important source of such poorly controlled promoters in the course of nonviral oncogenesis. □

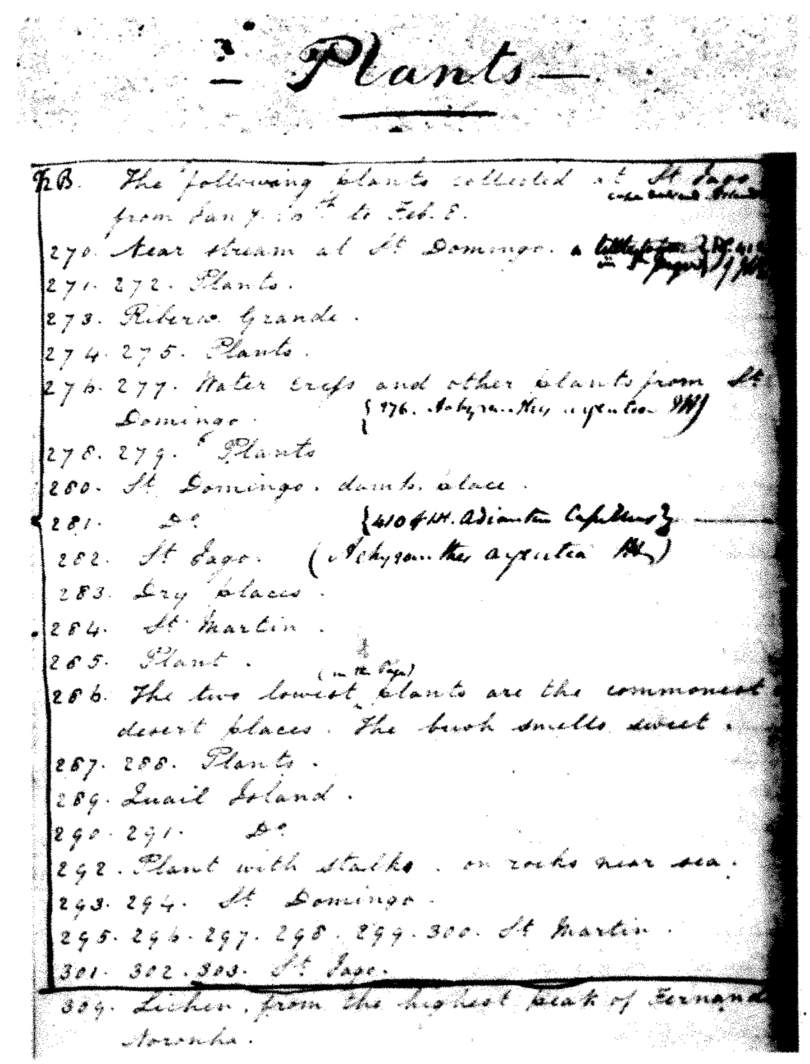
---

* Institute of Molecular Biology, Moscow.
†Institute of Molecular Genetics, Moscow.

0028-0836/81/190012-01$01.00

Darwin's missing notebooks come to light

from Duncan M. Porter



The cover legend written by Darwin and part of the first page of the Plant Notes written by Covington with notes added by Darwin and by Henslow.

IN early December 1980, Rita I'Ons, Assistant at the Cambridge University Herbarium, discovered two notebooks titled *C. Darwin's Plants from S. America* and *Darwin's S. American Plants* in a cupboard at the Herbarium. The notebooks, on the plants collected during the voyage of HMS *Beagle* (1831–1836) had been lost for more than a hundred years. Their long absence has caused many problems with the identification of Darwin's plants and has also prevented the botanical aspects of his voyage from being properly appreciated.

The notebooks are in the writing of J.S. Henslow, Professor of Botany at Cambridge and Darwin's mentor, who originally agreed to identify the plants. However, because of other commitments, Henslow was unable to identify more than a few, and the task fell to J.D. Hooker at Kew in the early 1840s. Henslow sent Darwin's notes to Hooker, and presumably they were later returned, but earlier searches at Cambridge and Kew failed to find them[1]. Hooker and others identified a number of Darwin's plants, but the majority remain unclassified to this day. Tied with string into the back of the second notebook are 11 pages of notes titled 'Plants'. They are in the handwriting of Syms Covington, Darwin's servant and amanuensis on the *Beagle* and until 1839. A few notes and corrections have been added in Darwin's hand, as have a few notes by Henslow.

Notes on the plants vary in length from the terse 'Plant' to a page-long discussion on the Wild Potato of Chile's Chonos Archipelago (*Solanum tuberosum* var. *guaytecarum*). The latter was used as the basis for the discussion in Darwin's *Journal of Researches*[2]. Likewise, one can trace other notes to passages in the *Journal*. Some, such as that on the Dwarf Beech of Tierra del Fuego (*Nothofagus antarctica*), make more interesting reading than when published: "Jemmy Button said. 'when leaves yellow, snow all go'. — Captain FitzRoy states that in April the leaves of the trees which grow on the lower parts of the hills turn colour; but not those high up. — I recollect having read a paper to show that in England warm Autumns hastened the falling of the leaves; that the process is a regular part of the vegetation: This fact would seem to show the same law." When published in the *Journal*, Button's quote was excised.

A passage in the Plant Notes describing the forests of Tierra del Fuego seems to anticipate the 'tangled bank' of the *Origin of Species:* "The extreme dampness of the climate favours the course luxuriance of the vegetation: the woods are an entangled mass where the dead and the living strive for mastery." In the last paragraph of the first edition of the *Origin*[3], Darwin writes: "It is interesting to contemplate an entangled bank, clothed with many plants of many kinds, with birds singing on the bushes, with various insects flitting about, and with worms crawling through the damp earth, and to reflect that these elaborately constructed forms, so different from each other, and dependent upon each other in so complex a manner, have all been produced by laws acting around us."

The Plant Notes follow the same format as the lists of notes on *Beagle* zoological specimens at the Cambridge University Library and the British Museum (Natural History). Each series of notes was compiled for the taxonomists who were to identify the various groups of organisms collected (plants, birds, fish, insects, mammals, molluscs, and reptiles and amphibians). As Darwin intended, the Plant Notes are proving quite useful in the identification of his *Beagle* specimens. ☐

Duncan M. Porter is Associate Professor of Botany at Virginia Polytechnic Institute and State University, Blacksburg, and Visiting Fellow at Clare Hall, University of Cambridge.

---

1. Porter, D. M. J. *J. Soc. Biblphy nat. Hist.* **9**, 515 (1980); *Bot. J. Linn. Soc., Lond.* **81**, 79 (1980).
2. Darwin, C. *Narrative of the Surveying Voyages of His Majesty's Ships Adventure and Beagle, Between the Years 1826 and 1836, Describing Their Examination of the Southern Shores of South America, and the Beagle's Circumnavigation of the Globe*, Vol. 3, *Journal and Remarks, 1832-1836* (Colburn, London, 1839); second edn: *Journal of Researches into the Natural History and Geology of the Countries Visited During the Voyage of H.M.S. Beagle Round the World, Under the Command of Captain FitzRoy, R. N.* (Murray, London, 1845).
3. Darwin, C. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*, 489 (Murray, London, 1859).

0028-0836/81/190013 01$01.00

The well-managed lawn

from Peter D. Moore

ONE of the most widespread vegetation types, and one which is certainly the most easily accessible to urban populations, is that highly managed piece of anthropogenic grassland, the domestic lawn. It has been estimated that the area of lawn in Britain is approximately 2,700 km^2 (NERC report, *Amenity Grasslands*; 1977), plus a further 1,100 km^2 of playing fields. It is surprising, therefore, that so little academic research has been conducted upon so amenable a habitat.

The work of Falk (*Ecology* **57**; 141, 1976) on his suburban lawn in California made it evident that the lawn is a very appropriate ecosystem for studies of the energetic type. Falk produced a very full energy budget for his lawn, in which net primary productivity was estimated at 4,468 kcal $\text{m}^{-2}\text{yr}^{-1}$ (1.05 kg m^{-2}). Of this 10.4 per cent was harvested by man as mowings, 2.9 per cent was consumed by invertebrate grazers and the remaining 86.7 per cent was considered to enter the decomposer system directly.

Falk's data also take into account the energy investment by man in the management and maintenance of the lawn, mainly in the form of fossil fuel used by the motor mower. This totals 2,029 kcal $\text{m}^{-2}\text{yr}^{-1}$. Thus if we were to consider the lawn in terms of harvest efficiency the outcome is low, with an output : input energy ratio of 0.22. The ratio obtained (Blaxter, *Nature* **252**; 531 1974) for the total energy efficiency of British agriculture is 0.4.

The British lawn is still neglected, but Falk has now provided us with more information about the lawns of the United States (*J. appl. Ecol.* **17**, 689; 1980), this time in the east, near Washington DC. Two lawns were selected for study, one received what was termed 'minimum maintenance', with no irrigation or fertilization and a fortnightly mowing. The other site was well maintained, being fertilized and limed once a year, mowed weekly and irrigated during the summer. Both lawns were mowed to a height of 5 cm using a rotary mower.

As one might expect, the undermanaged lawn was richer in species (22 species recorded in 0.7 ha) than was the highly managed one (11 species in 0.2 ha). What is surprising is the very similar values obtained for the total net primary productivity (below-ground production is included). For the neglected lawn this came to 1,669 $\text{kg m}^{-2}\text{yr}^{-1}$ (6,380 kcal) and for the well-managed lawn 1,649 $\text{kg m}^{-2}\text{yr}^{-1}$ (6,814 kcal). So all the extra management effort serves only to reduce diversity without increasing productivity. The amount removed in mowings is about 20 per cent of the total productivity, rather higher than that observed in California.

As Falk points out, the mowed lawn in temperate climes is a highly productive ecosystem, despite its low stature. The highest productivity of tropical grasslands (around 2,000 g $\text{m}^{-2}\text{yr}^{-1}$) and the productivity of maize or wheat in temperate environments (1,000–1,600 g $\text{m}^{-2}\text{yr}^{-1}$) are comparable to that of the suburban lawn. It is salutary to reflect, therefore, upon the energy harvest each year from such habitats in Britain alone. If these can match the productivity of their Washington counterparts (and anyone who has seen the lawns of Washington in August will have no hesitation in allowing this), then 332 kg m^{-2} (1,320 kcal m^{-2}) is harvested as mowings each year from our 3,800 km^2 of urban and suburban amenity grasslands. This amounts to 1.2×10^9 tons of dry matter, or 5×10^{12} kcal each year. Taking the original estimate of energy output : input ratio of 0.22, this implies that we expend a total of 2.27×10^{13} kcal on the harvesting of this herbage each year in Britain. Given the high energy investment in maintaining the British lawns and cricket pitches, the energy-from-biomass lobby must be tearing out its hair to see mowings being fed to the decomposer systems of the nations's compost heaps. □

Peter D. Moore is in the Department of Plant Sciences, University of London King's College.

---

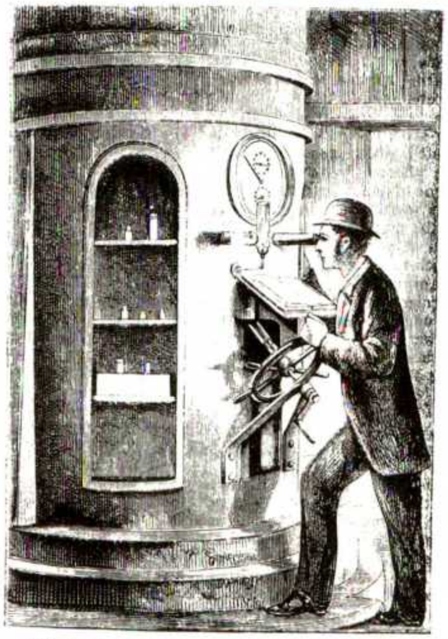
100 years ago

THE GREAT VIENNA TELESCOPE

THE political and social disturbances in Ireland have of late somewhat diverted attention from the literary and scientific work done in that country. Such work has nevertheless proceeded on its quiet way despite land agitation, failure of crops, or even commercial distress; and Ireland is to be congratulated on the completion of the fine 27-inch refracting telescope, designed and constructed by Mr. Howard Grubb of Dublin for the Imperial and Royal Observatory of Vienna. This telescope is the largest equatorial refractor at present in existence.

In the largest telescope until now in existence, the great 26-inch refractor at Washington, the oft-recurring operation of reading the circle involved the sending of an assistant with a hand-lamp to climb up some twenty feet of the instrument, and the vernier not being in a fixed place, but moving about with the telescope, the labour and difficulty of this operation are great. In the Vienna instrument Mr. Grubb has so contrived it that the observer sitting in his chair can read the right ascension and declination circles through one single reader, all being illuminated by one gas-lamp hung on the end of the declination axis.

Another wonderful convenience is enabling an assistant to easily turn and set the instrument to right ascension. This can be understood from the figure, which is from a photograph. In the case of an instrument of the size of the Vienna equatorial, the observer requires an elaborate stage or platform of variable height and position in order to reach the eye-end, and if he has to move to another object he must descend from his chair or platform and move *it* before he can move the instrument, and then he has not the facility of sighting objects, but must adopt a sort of tentative process, climbing up and down his stage and moving it and the telescope alternately. To avoid this labour and save time an arrangement is supplied by which the assistant from the ground floor can set the instrument in right ascension, and once set in one direction, the other (declination setting) can be readily managed by the observer.



We learn that M. Alphaud, the Director of Public Works in Paris, has in his hands the tender of Siemens and Co. for constructing an electrical railway from the Place de la Concorde to the interior of the Electrical Exhibition. M. Alphaud has given his adhesion to the request, which will be sent with his recommendation to the Commission of Sewers appointed by the Municipal Council, that when the Exhibition shall be closed, the railway will be kept running in the Champs Elysèes.

About twenty minutes to eleven on Monday night, owing to some accident at present unexplained, the electric lights on the Brush system, one of the three with which experiments are at present being made in the City, were suddenly extinguished, leaving a large portion of the City in total darkness. The area over which the Brush light has been placed extends from Blackfriars Bridge, up Ludgate Hill, to St. Paul's Churchyard, and down Cheapside as far as Queen Street and King Street. Fortunately the old gas-lamps remain in their places while the electric light experiments are being made, and orders were quickly given for these to be lighted. Every attempt was made by those in charge of the Brush light to restore the connection, and for a few minutes it seemed as though they had succeeded; but this only lasted for a very short time, and it was soon seen that something had gone hopelessly wrong.

from *Nature* **24**, 5 May, 11, 17, 1881

0028-0836/81/190014-01$01.00

Carbon in carbonaceous chondrites: carbyne or graphite?

from P. P. K. Smith

CARBON in carbonaceous chondrite meteorites has received much attention as a likely carrier of isotopically anomalous noble gases. Over the past year this has led to a difference of opinion concerning the structural form of the carbon in such meteorites. This problem was addressed by several contributors to the Twelfth Lunar and Planetary Science Conference, held in Houston, Texas, in March.

Carbonaceous material has been reported[1,2] as a carrier of part of the anomalous neon component, Ne-E, which is characterised by its extreme enrichment in the isotope ^{22}Ne. In 1969 Black and Pepin[3] argued that Ne-E is of presolar origin, having been trapped in interstellar dust grains that were subsequently incorporated into the solar nebula. A carbonaceous host phase also seems likely[1] for a presolar xenon component, s-process Xe, that was discovered by Srinivasan and Anders[4]. Thus if the carrier phases of these exotic noble gas components can be identified they may provide us with examples of interstellar dust grains.

Last year Whittaker and Watts of the Aerospace Corporation, in collaboration with Lewis and Anders of Chicago University, reported that 80 per cent or more of the carbon in the Allende carbonaceous chondrite is in the form of carbynes[5]. Carbynes are a relatively recently discovered series of polymorphs of elemental carbon, based on a chain structure with alternating single and triple bonds[6]. Whittaker *et al.* identified carbynes in Allende by means of electron diffraction, and also by the characteristic negative-ion mass spectrum that they produce in an ion probe. Carbynes were also detected in the Murchison carbonaceous chondrite, where they were found to be carriers of Ne-E and s-process Xe.

A different view of the structure of the Allende carbon has been presented by Smith and Buseck of Arizona State University, who used high-resolution transmission electron microscopy (HRTEM) to obtain images of the microstructure of the carbon[7,8]. Such images display prominent crystal lattice fringes with spacings between 3.4 and 3.9Å, corresponding to the interplanar spacings of poorly crystalline graphite. The individual graphite crystallites are only ~ 100Å in size, so that electron diffraction patterns obtained from areas 0.5 μm across (the smallest area that can be isolated by conventional selected-area electron diffraction) consist of diffuse rings. This may be contrasted with the sharp, single-crystal diffraction patterns that Whittaker *et al.* obtained from carbyne crystals which were up to 0.75 μm in size.

At the recent conference Smith and Buseck reported an attempt to resolve this discrepancy[9]. The Chicago/Aerospace group had provided a portion of their nearly pure carbon separate (BK) from Allende, together with an electron microscope mount that they had prepared from this material. Smith and Buseck located several grains on this electron microscope mount that gave carbyne-like diffraction patterns. However, chemical analysis of these grains by means of X-ray microanalysis demonstrated that they were all sheet silicates rather than elemental carbon. Examination of an electron microscope mount of BK that Smith and Buseck prepared themselves revealed only graphitic carbon. Thus Smith and Buseck inferred that some, and possibly all, of the diffraction patterns that Whittaker *et al.* obtained were actually due to contamination of their sample by sheet silicates. This would appear to account for one unexplained feature of the Chicago study: although BK was considered to be 80 per cent carbyne, only five carbyne crystals were found in it by electron diffraction.

Of course these observations do not have any direct bearing on the ion-probe results of Whittaker *et al.*, and in discussion of the Arizona study Anders emphasised the striking predominance of even-numbered carbon species (C_2, C_4, etc.) that they had found in the negative-ion spectrum of the Allende material. This pattern is believed to reflect the tendency of the . . . C–C≡C–C . . carbyne chain to cleave at the single bonds rather than at the triple bonds[10]. This feature is not in fact unique to carbynes — a predominance of C_{2n} species has been described previously for mass spectra obtained from both graphitic[11] and amorphous[12] carbon — but Whittaker *et al.* used reference spectra from graphite and from known carbynes to determine the carbyne content of their Allende carbon separate. The term graphitic carbon embraces a wide variety of microstructural forms, ranging from macroscopic single crystals to the extremely fine-grained tangled assemblage of ribbon-shaped crystallites that constitutes 'glassy' carbon[13]. Further work is perhaps needed to determine whether these varieties give rise to differing negative-ion spectra.

Electron diffraction observations of Allende carbon reported by Lumpkin[14] are consistent with the HRTEM observations of graphitic carbon. Lumpkin did find a small fraction (~ 3 per cent) of grains that gave carbyne-like diffraction patterns, but the chemistry of these grains has not yet been verified by microanalysis.

Dziczkaniec *et al.* described an investigation of the trapping of noble gases during the synthesis of carbonaceous material in a plasma discharge, an experiment designed to model the conditions under which the carbonaceous carrier phases in meteorites may have formed[15]. Although at an early stage, this study has shown that isotopic fractionation takes place during the trapping of xenon. It will be interesting to see whether the release temperatures for noble gases in this synthetic carbon follow a pattern similar to that found for meteoritic carbons. Interestingly enough, Dziczkaniec *et al.* found some crystalline grains amongst their run products that were very tentatively identified as carbynes. If this identification is confirmed, then this study would provide a new synthesis technique for carbynes, which have usually been prepared by laser heating of graphite.

Most recent studies of carbonaceous material in meteorites have centred on carbon-rich residues obtained by dissolving the bulk of the meteorite in concentrated acids, thereby destroying all information on the petrographical distribution of the carbon. An exception to this was the discovery by Scott *et al.* of aggregates of graphite and magnetite (Fe_3O_4) occurring as millimetre-size inclusions in several ordinary chondrite metoerites, and also in some more primitive unequilibrated ordinary chondrites[16]. The presence of this graphite–magnetite mixture in such primitive meteorites, which are thought not to have undergone appreciable heating in their asteroid parent bodies, indicates that graphite and magnetite may have been significant constituents of the early solar nebula. Textural relationships in these meteorites suggest that much of the

---

P. P. K. Smith is a Postdoctoral Research Associate in the Department of Geology, Arizona State University.

1. Lewis, R. S., Alaerts, L., Matsuda, J. & Anders, E. *Astrophys. J.* **234,** L165 (1979).
2. Eberhardt, P., Jungck, M. H. A., Meier, F. O. & Niederer, F. *Astrophys. J.* **234,** L169 (1979).
3. Black, D. C. & Pepin, R. O. *Earth planet. Sci. Lett.* **6,** 395 (1969).
4. Srinivasan, B. & Anders, E. *Science* **201,** 51 (1978).
5. Whittaker, A. G. *et al. Science* **209,** 1512 (1980).
6. Whittaker, A. G. *Science* **200,** 763 (1978).
7. Smith, P. P. K. & Buseck, P. R. *Meteoritics* 15, 368 (1980).
8. Smith, P. P. K. & Buseck, P. R. *Science* (in the press).
9. Smith P. P. K. & Buseck, P. R. *Lunar planet. Sci.* **12,** 1017 (1981).
10. Stuckey, W. K. & Whittaker, A. G. *10th Carbon Conference*, Bethlehem, Pa. (1971).
11. Honig, R. E. *J. chem. Phys.* **22,** 126 (1954).
12. Fürstenau, N., Hillenkamp, F. & Nitsche, R. *Int. J. Mass Spectrom Ion Phys.* **31,** 85 (1979).
13. Jenkins, G. M., Kawamura, K. & Ban, L. L. *Proc. R. Soc. Lond.* **A327,** 501, (1972).
14. Lumpkin, G. R. *Lunar planet. Sci.* **12,** 631 (1981).
15. Dziczkaniec, M., Lumpkin, G. R., Donohoe, K. & Chang, S. *Lunar planet. Sci.* **12,** 246 (1981).
16. Scott, E. R. D. *et al. Lunar planet. Sci.* **12,** 955 (1981).
17. Niederer, F. & Eberhardt, P. *Meteoritics* **12,** 327 (1977).

0028-0836/81/190015-02$01.00

metallic iron present in them formed by the reduction of magnetite by carbon.

A previous report[17] of presolar Ne-E in a carbon-rich residue from one of the ordinary chondrites studied by Scott *et al.* led them to suspect that the graphite–magnetite inclusions might be the carrier, and hence be of presolar origin. Unfortunately the story is not that simple — they did not find any Ne-E in the graphite–magnetite intergrowths. Thus the Ne-E must either be contained in a different carbonaceous component, or in another minor phase in the residue. The identification of the carriers of such minor gas components is fraught with difficulties, but is providing us with valuable clues to the early history of the Solar System. □

Impact cratering

from Richard A.F. Grieve

The relative importance of cratering in planetary evolution was evidenced at the Twelfth Lunar and Planetary Science Conference* by numerous papers on impact craters and discussion, in sessions dealing with particular terrestrial planets and the icy satellites of Jupiter and Saturn, of their effects. The diversity of research reported in the sessions dealing directly with impact processes reflected the random nature of cratering with respect to specific geological environments and the involvement of many sub-disciplines of the physical science in providing information on cratering processes. As has been apparent in recent years, there was an increasing combination of theoretical and observational results, with each approach serving to constrain the other.

With respect to the evolution of the jovian and saturnian satellites, the results of impact experiments showed that the mechanical properties of ice differ from those of rock, even at the low temperatures of 70–80K existing on these bodies. In particular, the specific energy required to disrupt an icy body by impact appears to be ~ 20 per cent less than for a silicate body. Other impact experiments illustrated that extensive fracturing occurs around craters in ice but craters in ice-saturated sand targets exhibit more limited fracturing and are more like those in competent rock. It was noted, however, that ejection patterns appear to be similar for craters in icy and silicate targets.

Computational simulations derived from the solution of continuum mechanics equations continued to provide important insights into cratering processes. For example, they indicated that excavation dynamics and the physical state and radial distribution of ejecta are strong functions of impact velocity. Calculations also supported the interpretation that bowl-shaped cavities are produced by excavation plus the displacement of the cavity floor by impact-induced flow fields. Further confirmation of these displacements was proved by an analysis of the Cactus nuclear event. Similar conclusions on the formation of bowl-shaped craters were reached in a combined observational and theoretical treatment of the Brent Crater, Canada, which also provided an estimate of the initial impact conditions: a 113-m radius chondritic projectile impacting at ~ 27 km s^{-1}.

The important question of the detailed particle flow fields created in the target rocks by impact was addressed in a comparison of flow fields derived from continuum mechanics calculations and the 'Maxwell-Z' model. It was clear that this empirically derived model remains a powerful analytical tool. Similarities between the calculated flow fields for Meteor Crater, Arizona, and those of the Middle Gust III explosive event also suggested that the parametric codes derived from the latter may be applicable to impact events. However, the energy coupling differences suggest that explosive events are imperfect analogues to impact events in terms of scaling crater dimensions with energy.

It is known that the initial transient cavities formed by impact events generally modify rapidly to more stable final forms. In bowl-shaped craters, this modification is thought to occur by failure of the transient cavity walls. It was proposed at the meeting, however, that cavity modification may result from the centripetal motion of the entire rim area, in a manner similar to that suggested for the formation of larger, relatively shallow complex impact structures. The formation of these complex structures was addressed in a phenomenological model, ascribing central peak and ring formation to oscillatory motions of the transient cavity floor. These oscillatory motions require essentially plastic behaviour and viscosity of 10^9 P for the target rocks. Such viscosities may result from the effective fluidization of fractured target by an impact-generated acoustic field. Some support for the importance of two parameters, rock strength and gravity, in the phenomenological model may be found in the report that the 20-km Haughton impact structure on Devon Island, Canada, has a multi-ring form. Haughton was formed in sediments and its relatively small dimensions for a ring structure scale with gravity and rock strength when compared with similar crater forms on the Moon and in crystalline targets on Earth. Doubts were raised, however, over the universal concept of deep initial transient cavities by the results of impact experiments at high gravitational accelerations, which suggested that relatively shallow cavities may be produced directly without modification.

Several papers dealt with the mineralogical or shock metamorphic effects produced by impacts in various target media. Shock experiments in limestone indicated that incipient vapourization of carbonate occurs at pressures as low as 350 kbar. An analysis of effects in tectosilicates illustrated that planar deformation features are produced during shock compression. It was also reported that the reduction in refractive index of shocked quartz is due to lattice disordering. A high concentration of lattice defects, particularly non-bridging oxygens, is also the cause of the conversion of plagioclase to the solid-state glassy phase known as maskelynite at shock pressures of 300–350kbar.

The non-isochemical alteration of impact melt rocks was reported, as was evidence for extensive chemical exchange between feldspar clasts and enclosing impact melt matrix. This chemical exchange may account for some of the disturbances seen in the Rb–Sr systematics of melt rocks. ^{39}Ar–^{40}Ar systematics of shocked samples from the Ries, Germany, were also described and it appears that up to 99 per cent of radiogenic ^{40}Ar may be lost at shock pressures of 300 kbar or more. This observation should aid in the interpretation of Ar ages obtained on lunar impact breccias.

A report of siderophile abundances in the melt rocks at Strangways suggested that this 20–25 - km impact structure in Australia was formed by an olivine-rich achondrite. Siderophile enrichments were also discussed with reference to the Cretaceous–Tertiary boundary layer and extinctions. It was noted that several important geochemical problems with respect to the various sampled sections need to be resolved. Theoretical calculations were presented for the postulated extinction-causing impact and indicated that significant atmospheric heating could only be achieved indirectly through energy transfer from ejecta. The interpretation that the formation of the 20-km lunar crater Giordano Bruno is recorded in the Gervase of Canterbury chronicles as occurring on 19 June 1178 was reviewed. It appears that only an estimated 0.1 per cent of the ejecta might reach the Earth by direct trajectories and that the analytical problems associated with recognizing such material are formidable. □

Richard A.F. Grieve is Visiting Professor at the Department of Geological Sciences, Brown University, Rhode Island.

---

*The Twelfth Lunar and Planetary Science Conference was held in Houston, Texas in March.

0028-0836/81/190016-01$01.00

ARTICLES

The structure of West Antarctica from geophysical studies

E. J. Jankowski & D. J. Drewry

Scott Polar Research Institute, University of Cambridge, Lensfield Road, Cambridge CB2 1ER, UK

The subglacial topographical and geological configurations of West Antarctica and the boundary zone between East and West Antarctica are poorly known, yet they are crucial to our understanding of the tectonic evolution of the Pacific–proto Atlantic margin of Gondwanaland. Preliminary results are presented of simultaneous airborne radio echo and magnetic sounding of $\sim 10^6$ km^2 of West Antarctica including detailed maps of ice sheet and bedrock surfaces. Magnetics, radio echo and other available geophysical data suggest interpretations of sub-glacial geology.

DURING the late 1950s and 1960s major US oversnow traverses in West Antarctica conducted altimetric, seismic, gravimetric and magnetic studies[1–4]. This work provided the first outline of ice sheet surface, subglacial topography and ice thickness. Many features were identified, including the principal ice divides, subglacial basins and mountainous blocks. Airborne magnetometry was also conducted, principally by the University of Wisconsin[5,6] the US Geological Survey[7] and US Project Magnet[8,9].

The oversnow traverses accomplished some 2,000 surface altitude determinations, ~250 seismic soundings and ~2,000 gravity observations in an area of $\sim 10^6$ km^2. Some interpolation and extrapolation was necessary for maps to be produced with a contour interval[10,11] of 500 m. Although the aeromagnetic data had high accuracy, they suffered from having no simultaneous ice thickness information and only dead reckoning navigation.

During the austral summers 1977–78 and 1978–79 the Scott Polar Research Institute (SPRI) undertook an airborne geophysical programme in West Antarctica which included radio echo sounding (by SPRI and Technical University of Denmark, TUD) and magnetometry (by Applied Physics Laboratory, APL, Johns' Hopkins University). The work was carried out in cooperation with the US National Science Foundation (NSF). In addition radio echo sounding (RES) alone was conducted during the austral summer 1974–75 in the area between Byrd Station and the Ross Ice Shelf[12] (see Fig. 1 for locations). The principal objective was to establish a grid for geophysical investigation of the complex structure of West Antarctica and its relationship to the East Antarctic craton. The available geophysical data combined with reconnaissance geology on scattered outcrops suggest West Antarctica possesses a complicated tectonic history[13,14].

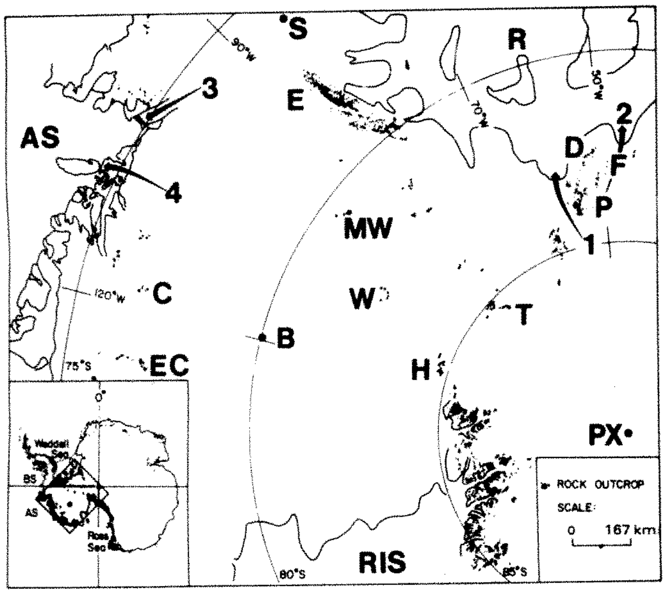


Fig. 1 Location map. AS, Amundsen Sea; BS, Bellingshausen Sea; R, Ronne-Filchner Ice Shelf; RIS, Ross Ice Shelf; B, Byrd Station; C, Crary Mountains; D, Dufek Massif; E, Ellsworth Mountains; EC, Executive Committee Range; F, Forrestal Range; H, Horlick Mountains; MW, Mount Woollard; P, Pensacola Mountains; PX, Pole Station; S, Siple Station; T, Thiel Mountains; W, Whitmore Mountains; 1, Foundation Ice Stream; 2, Support Force Glacier; 3, Pine Island Glacier; 4, Thwaites Glacier.

Equipment during both seasons comprised pulsed radar systems operating at centre frequencies of 60 and 300 MHz, designed and built by TUD in collaboration with SPRI[15,16]. Returned echoes were displayed in both *A*- and *Z*-modes and recorded on 35 mm film and dry-silver paper by a fibre-optic oscillograph. Equipment was mounted on pallets in a specially configured NSF C-130R Hercules aircraft, operated by US Navy Antarctic Development Squadron Six. An LTN-51 inertial system provided navigation data updated by photographic resections on known topographic points. Navigation and all other flight data were recorded on magnetic tape by an Airborne Research Data System (ARDS) constructed by US Naval Weapons Center, and operated since 1978 by APL. A proton precession magnetometer mounted in a tail stinger was installed by APL and operated by them simultaneously with RES. Coefficients compensating for the effect of the aircraft were determined by APL using a vector magnetometer and adopting the technique originally developed by Leliak[17–19]. The unreduced magnetic data, provided by the NSF, were then compensated, reduced and correlated with RES in Cambridge.

Aircraft pressure altitude was corrected to points of known surface elevation. The terrain clearance was determined from RES records and the aircrafts' radar altimeter. Navigation data were corrected for closure error, typically of the order of 3 km, and inflight fixes. Subsequent errors in altimetry from one flight to another were removed using a random walk technique. The mean error at crossover points before random walking was 46.5 m with a standard deviation of ±44 m. RES records were digitized at a spacing of between 1.5 and 1.8 km and concatenated with updated navigation and altimetry data. Maps were compiled to depict ice sheet surface and bedrock configuration. In the case of the ice-sheet surface some 70,000 RES data points were used. Additional elevations were obtained by using all data from oversnow traverses (refs 1–4, 20 and C. R. Bentley, personal communication). For the bedrock map we used ~56,000 RES determinations combined with bedrock elevations at some 250 seismic stations on the traverses. A contour interval of 100 m has been used for the ice surface and 250 m for

0028-0836/81/190017—05$01.00

bedrock. Shading has been applied to areas of bedrock below −1,000 m a.s.l. to assist with interpretation. No attempt was made to detail the contours in areas of outcrop.

Ice sheet surface

Although there is broad correspondance between Fig. 2 and previous maps based on oversnow traverse[20,21] or balloon data (tropical wind, energy conversion and reference level experiment, TWERLE)[22], the large amount of data now available from RES allows considerably better delineation of many features including the ice drainage basins of West Antarctica. A major ice divide runs from inland of the Transantarctic Mountains through a gap between the Horlick and Thiel Mountains towards the Ellsworth Mountains confirming previous studies. A further ice divide crosses from the Executive Committee Range (passing to the north-east of Byrd Station) to join the first divide in the vicinity of Mt Woollard (see Fig. 1 for locations). West Antarctica is thus effectively divided into a series of drainage basins feeding ice into the Ross Ice Shelf, Ronne–Filchner Ice Shelf and the Amundsen–Bellingshausen Seas. Drainage into the Ross Ice Shelf from West Antarctica has been fully described elsewhere[23].

Ice streams shown shaded in Fig. 1 have been delineated by the presence of surface crevassing detected on the RES records[23]. The Foundation Ice Stream, running along the west side of the Pensacola Mountains, is already known in general but additional detail shows it to be a much more significant feature of West Antarctic drainage. Another ice stream is located 250 km further to the north-west. Although a less pronounced feature with boundaries lacking strong definition, its crevassed, linear margins are, nevertheless, clearly visible on RES records and from the air. It extends inland only for a few hundred kilometres. The area between these two ice streams forms a peninsula and remains one of the least known margins of the Ronne–Filchner Ice Shelf. Elsewhere LANDSAT satellite imagery has allowed the grounding line and ice rises to be delimited with considerable accuracy[24,25]. RES records in this area reveal an undisputed ice–water interface indicating ice shelf. This is supported by ice thickness–surface elevation relationships. Further inland characteristic ice–rock reflections showing rapid fluctuations in instantaneous returned power are observed. Between the ice streams, however, a curious 'intermediate' zone has also been detected where received power is high and the echo return is smooth over considerable distances (tens of kilometres), yet ice thicknesses and surface elevations do not demonstrate hydrostatic equilibrium. Furthermore, high-resolution photographs from the Soviet 'Meteor' weather satellite show no clear grounding line shadow as exhibited elsewhere

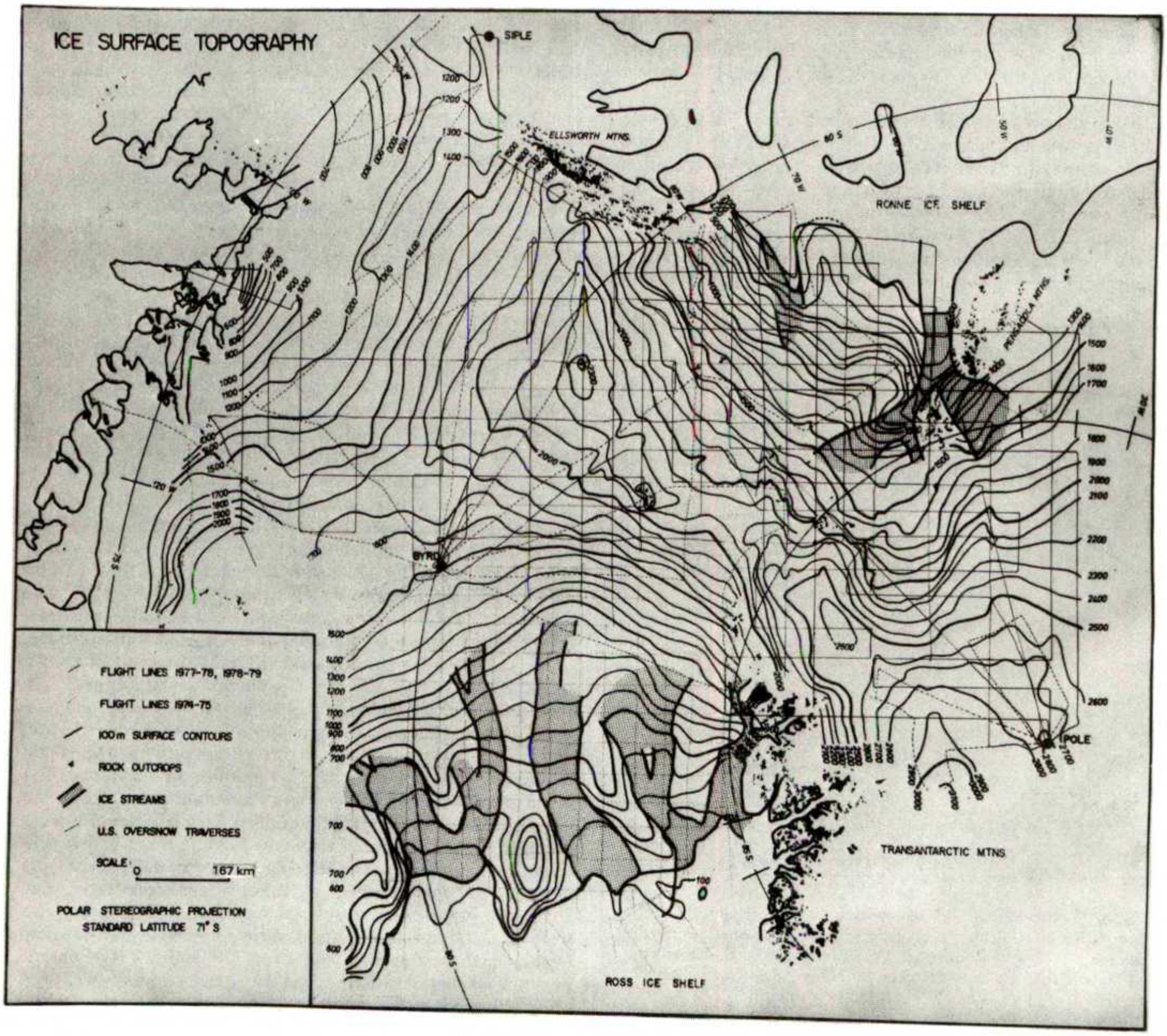


Fig. 2 Ice surface topography.

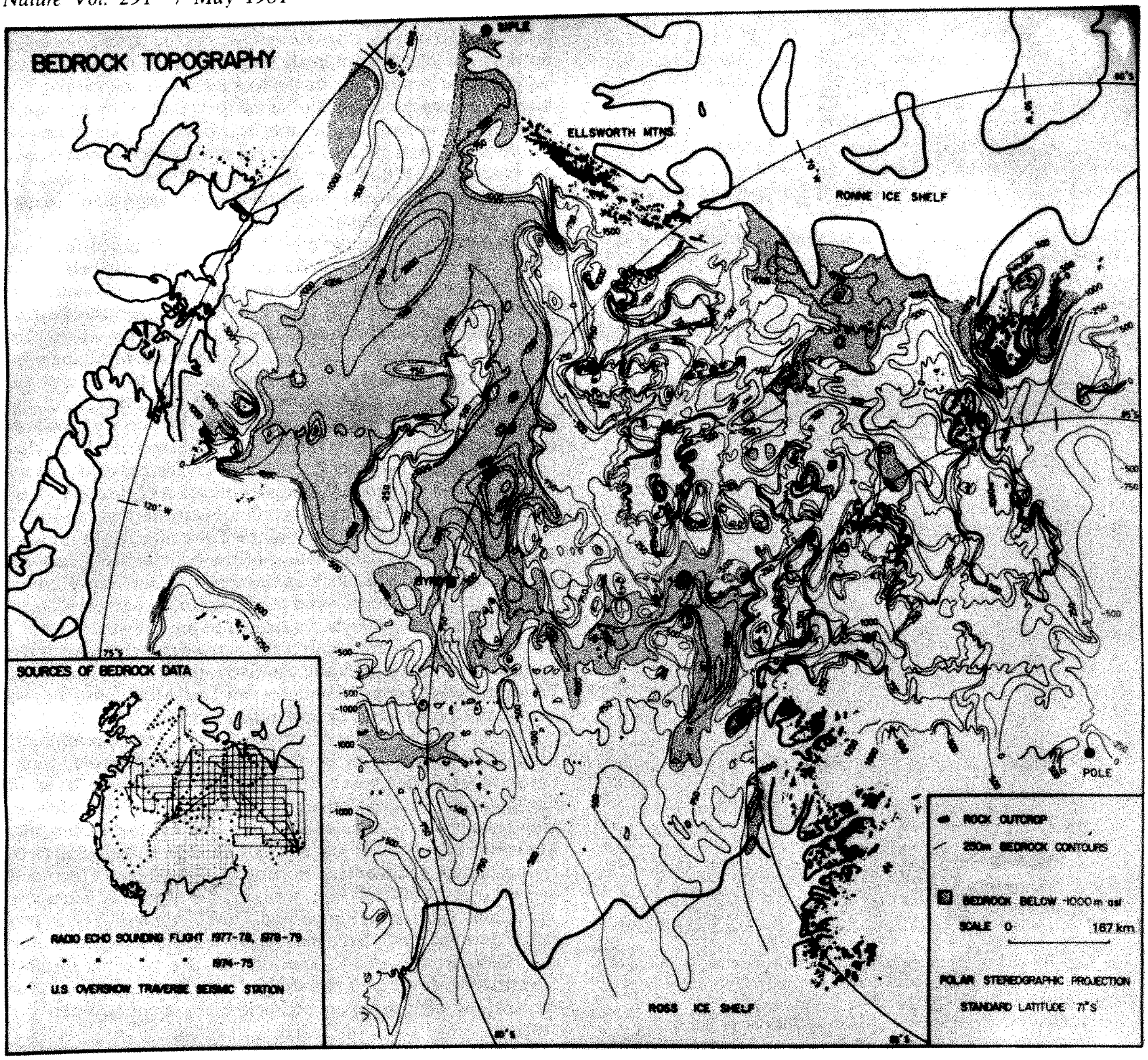


Fig. 3 Bedrock topography.

on the Ronne–Filchner Ice Shelf. The lack of relief–reflectivity further suggests that the area may possess an unusual and diffuse boundary[26]. We explain the 'intermediate' zone as ice shelf resting on sea floor composed of soft water-saturated sediments—a similar phenomenon to that observed by Soviet glaciologists on the Lazarev ice shelf in 1975[27].

Ice drains into the Amundsen Sea principally through the Thwaites and Pine Island glaciers. Unlike the ice streams flowing into the Ross and Ronne–Filchner Ice shelves these major outlets are not protected by large ice shelves and unless they flow over high bedrock sills may be in a finely balanced state of equilibrium[28] or may not even be in equilibrium at all[37]. RES suggests that Thwaites glacier does not have a high bedrock sill. There is, however, no pronounced erosional channel associated with the glacier which suggests that it may be a transient phenomenon. No data were collected on the Pine Island Glacier.

Bedrock configuration

The map presented in Fig. 3 is the first detailed and systematic study of bedrock using RES in West Antarctica. There is again a broad correspondance between major topographical features and previous maps based on seismic soundings and interpretation of free air gravity anomalies[10,11]. The location of all seismic stations used in this survey are shown in Fig. 3. Compared with previous work, however, the most striking feature of the map (and profiles) is the considerable improvement in topographical detail.

The area mapped can be divided into several parts some of which had been known from earlier work, and are briefly outlined here. (1) At the head of the Ross Ice Shelf the terrain is very smooth with subdued ridges and troughs marking the position of the ice streams draining into the ice shelf. This area is terminated to the north-east by (2) an extensive subglacial basin (Byrd Subglacial Basin). Insufficient data may account for its apparent smoothness. (3) A completely new feature is the sinuous ridge running across the centre of the basin. Near Byrd Station it is prominent on RES records, standing 1,000 m above the floor of the basin. (4) At the northern margin of the basin Cenozoic volcanic rocks crop out in the Crary Mountains and Executive Committee Range. RES records show these areas of outcrop to be steep sided and separated by deep narrow troughs. (5) A major highland region extends southwards from the exposed Ellsworth Mountains. The terrain is extremely rugged, possessing a relief amplitude of 2 km, a feature not previously apparent nor emphasized on the map due to spacing of flight lines and the contour interval. There is a very well defined boundary between this highland area and the Byrd Subglacial Basin. The Whitmore Mountains form a steep-sided upstanding

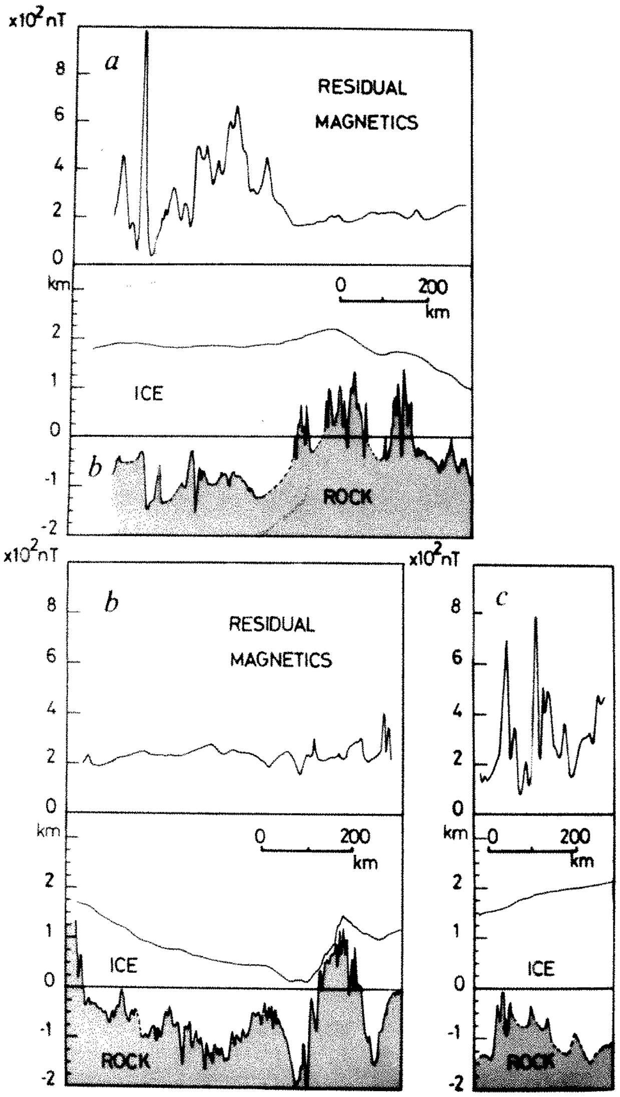


Fig. 4 RES and magnetic profiles in West Antarctica. *a*, Boundary between Byrd Subglacial Basin and Ellsworth Mountains block. *b*, Bedrock low at head of Ronne–Filchner Ice Shelf. *c*, Sinuous ridge cutting across Byrd Subglacial Basin.

massif separated from the Ellsworth Mountains block by a narrow trough. Other scattered outcrops to the south of the Ellsworth Mountains appear to be part of the same mountainous region. (6) The Transantarctic Mountains constitute a separate highland area and form a semi-continuous chain across Antarctica. Two deep valleys almost sever the Horlick Mountains from the rest of the Transantarctic Mountains, whereas the main part of the Pensacola Mountains is bounded on two sides by very deep narrow troughs occupied by Foundation Ice Stream and Support Force Glacier. At the northern end of the Pensacola Mountains there is a small depression between the exposed Forrestal Range and Dufek Massif. (7) At the head of the Ronne–Filchner Ice Shelf bedrock forms a deep depression which widens towards the ice shelf.

Geological interpretation

Bedrock at the head of the Ross Ice Shelf has been fully discussed elsewhere[12] where subdued topography and magnetic signatures are attributed to a substantial thickness of late Mesozoic–Cenozoic sediments extending the Ross Sea sedimentary basin beneath the grounded ice sheet. Further inland there is a significant change in magnetic character towards the volcanic outcrops in Northern Byrd Land and the Byrd Subglacial Basin, as noted in previous surveys[1,4–6]. Large anomalies (~1,000 nT) are here associated with outcrops and subglacial features. The sinuous ridge running across the basin possesses a pronounced magnetic anomaly (Fig. 4*c*): modelling in progress suggests that this feature may also have a volcanic origin. The magnetic character of the floor of the Byrd Basin shows a pattern of irregular anomalies. Seismic refraction studies from the centre of the basin indicate ~2 km of strata with P-wave velocity 4.1 km s^{-1} (ref. 29). These results suggest that the basin may constitute a sequence of interbedded sedimentary and volcanic strata overlying basement.

The pronounced topographical boundary between the Byrd Basin and rugged Ellsworth Mountains block is also marked by an abrupt change in pattern of magnetics from high-amplitude anomalies associated with the Basin to a magnetically quieter field with only isolated anomalies of wavelength ~40–50 km (Fig. 4*a*). The subdued magnetics are consistent with subglacial extension of a geological sequence similar to the exposed Ellsworth Mountains (Palaeozoic strata). This margin of the Ellsworth block becomes less well defined further west towards the Whitmore Mountains (south of Byrd Station); it is likely that sedimentation since the late Mesozoic has masked deeper structures and inhibits accurate identification of this boundary. A series of subglacial basins continues to the west and then trends south towards the front of the Transantarctic Mountains and may mark the continuation of the edge of the Ellsworth block. The Ellsworth block as described by the above boundaries includes an area around the Whitmore Mountains. These are thought to be geologically related to the Ellsworth chain[30], but may form a separate unit if the narrow trough between the two is structurally significant. Similarly, there is no clear break between the Whitmore Mountains and Thiel Mountains, the two being linked by a series of upstanding blocks.

Further south, magnetic profiles over the Transantarctic Mountains (except at the northern end of the Pensacola Mountains) suggest the presence of a significantly thick layer of non-magnetic rocks—possibly corresponding to Beacon Supergroup strata and older Precambrian metasediments. Modelling of anomalies just north of Thiel Mountains indicates a substantial sedimentary section[26]. The largest recorded anomalies (>1,000 nT) are seen in the northern Pensacola Mountains (Forrestal Range and Dufek Massif). These two ranges form part of a large mafic intrusion which has been fully described elsewhere[7,31,32]. The lopolith-like synclinal form of the intrusion, developed as a result of weak folding, accounts for the shallow depression in bedrock topography between the ranges[7].

The distinctive bedrock low at the head of the Ronne–Filchner Ice Shelf lies between the Transantarctic and Ellsworth Mountains, widening into the Weddell Sea. Hochstein calculated depth to MOHO at 83° S 70° W on the eastern edge of this basin, reporting thin crust of 24 km thickness[33]. Previous work has suggested that the Ellsworth Mountains and any associated subglacial block are out of place in West Antarctica based on their stratigraphy and structural trends[14,34–36]. The thin crust lying beneath the bedrock low may thus represent crust rifted during the displacement of the Ellsworth block. Magnetic records from this basin area show no significant anomalies and confirm earlier suggestions of a substantial section of sedimentary rock in the area[7] (Fig. 4*b*).

The present studies are being extended to provide quantitative estimates of rock magnetizations, depth of sedimentary cover, a detailed analysis of the structure and areal extent of the Dufek Massif (in collaboration with US Geological Survey), and to reconsider the position of West Antarctica in Gondwanaland.

This work was supported by NERC grant GR3/2291 and the NSF Division of Polar Programs. C. R. Bentley provided unpublished data from the Wisconsin traverse. We thank the NSF for aeromagnetic data, Dr J. C. Behrendt and the US Geological Survey Branch of Regional Geophysics for assistance in their reduction, and Drs G. de Q. Robin and D. H. Elliot for helpful discussions. TWERLE data were made available to SPRI by kind permission of N. Levanon and were plotted and contoured at SPRI for comparison with RES.

Received 31 October 1980; accepted 11 March 1981.

1. Bentley, C. R. & Ostenso, N. A. *Geophys. J.* **18**, 882–912 (1964).
2. Behrendt, J. C. *Geophys. Polar Res. Centre Rep.* 64–1 (University of Wisconsin, 1964).
3. Robinson, E. S. thesis, Univ. Wisconsin (1964).
4. Bentley, C. R. & Chang, F.-K. *Ant. Res. Ser.* **16**, 1–38 (1971).
5. Behrendt, J. C. & Wold, R. J. *J. geophys. Res.* **68**, 1145–1153 (1963).
6. Behrendt, J. C. *Science* **144**, 993–994 (1964).
7. Behrendt, J. C., Henderson, J. R., Meister, L. & Rambo, W. L. *US Geol. Surv. Prof. Pap.* 884 (1974).
8. *US Naval Oceanogr. Office, Spec. Publ.* 66 (1963).
9. *US Naval Oceanogr. Office, Spec. Publ.* 66, Suppl. 1 (1965).
10. Heezen, B. C., Tharp, M. & Bentley, C. R. *Am. geogr. Soc. N.Y.*, Folio 16 (Antarctic Map Folio Series, 1971).
11. Korotkevich, Ye. S., Koblents, Ya. P. & Kosenko, N. G. *Karta Korennogo Rel'yefa Antarktidy, 1:10000000* (Soyuzmorniiproyekt, Arctic and Antarctic Institute, Leningrad, 1977).
12. Rose, K. E. thesis, Univ. Cambirdge (1978).
13. Craddock, C. in *Antarctic Geoscience* (ed. Craddock, C.) (in the press).
14. Elliot, D. H. *Am. J. Sci.* **274A**, 45–106 (1975).
15. Skou, N. & Sondergaard, F. *TUD, Lab. Electromagn. Theor. Rep.* 169 (1976).
16. Christensen, E. *TUD, Lab. Electromagn. Theor. Rep.* **93** (1970).
17. Leliak, P. *IRE Trans Aeron. Nav. Electron.* September (1961).
18. Finkel, A. APL/S2R-78-194 (1978).
19. Hickerson, R. L. APL/S2R-79-145 (1979).
20. *Am. geogr. Soc. N.Y. 1:5000000, Sheet 13, Antarctica* (1970).
21. *Ministerstvo Morskogo Flota CCCP Karta Antarcktida* (1971).
22. Levanon, N., Julian, P. R. & Suomi, V. E. *Nature* **268**, 514–516 (1977).
23. Rose, K. E. *J. Glaciol.* **24**, 63–75 (1979).
24. Swithinbank, C. W. M., Doake, C. S. M. & Crabtree, R. D. *Polar Rec.* **18**, 295–299 (1976).
25. Crabtree, R. D. & Doake, C. S. M. *Polar Rec.* **20**, 31–37 (1980).
26. Drewry, D. J., Meldrum, D. T. & Jankowski, E. *Polar Rec.* **20**, 45–47 (1980).
27. Korotkevich, Ye S., Savtyugin, L. M. & Morev, V. A. *Inf. Byull. sov. Antarkt. Eksped.* No. 98, 45–52 (1978).
28. Thomas, R. H., Sanderson, T. J. O. & Rose, K. E. *Nature* **277**, 355–358 (1979).
29. Bentley, C. R. & Clough J. W. in *Antarctic Geology and Geophysics* (ed. Adie, R. J.) 683–691 (Oslo Universitetsforlaget, 1972).
30. Craddock, C. & Webers, G. F. in *Antarctic Geoscience* (ed. Craddock, C.) (in the press).
31. Ford, A. B. *US geol. Surv. Bull.* 1405–D (1976).
32. Behrendt, J. C., Drewry, D. J., Jankowski, E. & Grim, M. S. *Science* **209**, 1014–1017 (1980).
33. Bentley, C. R. *Tectonophysics* **20**, 229–240 (1973).
34. Schopf, J. M. *Science* **164**, 63–66 (1969).
35. Stump, E. in *Implications of Continental Drift to the Earth Sciences* Vol. 2 (eds Tarling, D. H. & Runcorn, S. K.) 909–924 (Academic, London, 1973).
36. Clarkson, P. D. & Brook, M. *Nature* **265**, 615–616 (1977).
37. Stuiver, M., Denton, G. H., Hughes, T. & Fastook, J. L. in *The Last Great Ice Sheets* (eds Denton, G. H. & Hughes, T.) 319–435 (Wiley, New York, 1981).

Parental liquids of the Skaergaard intrusion cumulates

C. E. Ford

Grant Institute of Geology, University of Edinburgh, West Mains Road, Edinburgh EH9 3JW, UK

Petrogenetic relationships between compositions advocated as possible liquids parental to the Skaergaard layered series cumulates have been investigated using equations for the prediction of liquidus temperature and liquid path followed during fractional crystallization. Given a suitably defined sequence of primocryst mineral assemblages, a variation of the fractionation scheme provides a new method for the calculation of the parent liquid composition.

IGNEOUS diversity may result from a combination of processes such as partial melting, magma mixing, contamination and fractional crystallization. However, the relative role of each process can only be accurately assessed if its effect can be separately predicted. This article considers the prediction of the compositions of evolving magmatic liquids, and complementary cumulus mineral assemblages, during fractional crystallization. The results are applied to the Skaergaard intrusion of east Greenland[1] which contains features expected of a cumulate sequence produced by fractional crystallization and also presents petrological problems.

Within an envelope of chilled marginal gabbros crystallization of the Skaergaard intrusion progressed inwards simultaneously from the roof, walls and unexposed floor where the upper and marginal border groups and the at least 2.5 km thick layered series were formed respectively. Segregation of cumulus minerals from liquid, with some sorting of minerals into layers, led to progressive evolution of the residual liquid composition and a complementary evolution of successive cumulus mineral assemblages. The early formed cumulates contain calcic plagioclases(An77/72) and magnesian olivines(Fo81) while the latest formed cumulates contain more sodic plagioclase(An30) and ferromagnesian minerals with extreme iron enrichment(Fo0). The points of appearance or disappearance of cumulus phases in the fractionation sequence have been used to subdivide the intrusion into zones. Particularly important are the point of appearance of cumulus clinopyroxene marking the base of lower-zone-b (LZb) and the points of disappearance, and subsequent re-appearance, of cumulus olivine marking the base and top of the middle-zone (MZ).

Composition of the Skaergaard parent liquid

The composition of the Skaergaard parent liquid is uncertain. Samples of the chilled marginal gabbro are heterogeneous[2] and may have been affected by wall rock contamination. Chill samples, which have been considered as possible parent liquids, cover a significant range of compositions. A parent liquid of the composition of the comparatively primitive chill EG4507 (Table 1) would require a 'hidden zone' of cumulates comprising some 70–80 wt% of the intrusion, based on estimates of the volume and composition of the exposed rocks and the proportion of trapped inter-cumulus liquid[1,3]. A hidden zone of this extent is consistent with calculations[4], based solely on arguments concerning olivine fractionation, which have suggested that 72

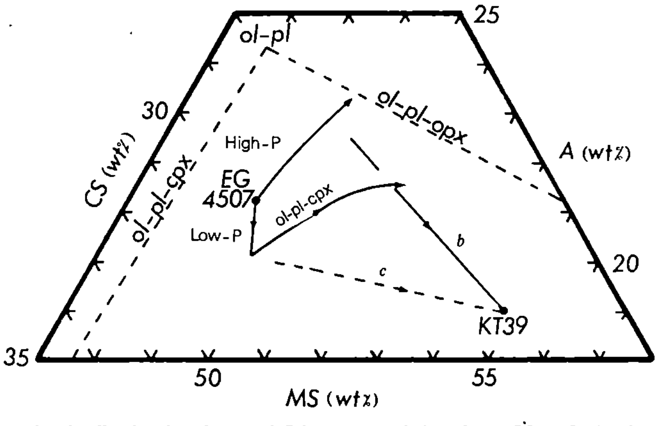


Fig. 1 Projection from M_2S into part of the plane $\dot{C}S$–MS–A of EG4507, KT39 and derived liquid compositions. At high pressure, crystallization of ol+pl+cpx from EG4507 drives the liquid composition away from the ol–pl–cpx tie-plane. At low pressure, crystallization of ol+pl precedes crystallization of ol+pl+cpx. Curve *b* represents negative fractional crystallization of ol+pl from KT39. This curve does not intersect those representing fractional crystallization of EG4507 at either low or high pressure—at the apparent points of intersection the mineral compositions are significantly different. Curve *c* is parallel to the liquid path followed in many cases of equilibrium crystallization of ol+pl+cpx.

~~R516~~ R417

0028-0836/81/190021—05$01.00

Table 1 Composition of Skaergaard chills and three possible cumulus enriched bulk compositions

	KT39	EG4507	B1	B2	B3
SiO_2	50.10	48.08	44.10	48.01	46.85
TiO_2	2.64	1.17	0.59	0.59	0.59
AlO_3	13.40	17.22	8.61	2.33	20.40
Fe_2O_3	1.77	1.32	0.66	0.66	0.66
FeO	11.81	8.44	11.12	4.22	6.26
MnO	0.20	0.16	0.08	0.08	0.08
MgO	6.54	8.62	27.34	4.31	11.10
CaO	10.09	11.38	5.69	13.89	11.48
Na_2O	2.42	2.37	1.19	2.28	1.96
K_2O	0.57	0.25	0.13	0.13	0.13
P_2O_5	0.22	0.10	0.05	0.05	0.05
H_2O^+	—	1.01	0.51	0.51	0.51
Total	99.76	100.12	100.07	100.06	100.07
Predicted liquidus mineral compositions			Predicted solidus mineral compositions		
O1	Fo 76(74)*	Fo 86	Fo 82	Fo 65	Fo 76
P1	An 73(67)*	An 81	An 62	An 74	An 73

KT39 from ref. 6, EG4507 from ref. 1. B1 = EG4507 + 50% olivine (Fo 86). B2 = EG4507 + 50% plagioclase (An 81). B3 = EG4507 + 15% olivine (Fo 86) + 35% plagioclase (An 81).

* Composition stated to be present in experimental charge.

cation% of fractional crystallization of chill EG4507 would be necessary to produce olivines of the composition observed at the exposed base of the layered series. However, geophysical evidence precludes the existence of such an extensive hidden zone immediately beneath the exposed layered series[5]. A parent liquid of the composition of the much more evolved, but less ubiquitous chill KT39 could perhaps produce the layered series with a minimal hidden zone[6]. The relationships between the various Skaergaard chill compositions and the layered series are thus not easily explained.

Evolving Skaergaard liquids and coexisting primocrysts

To calculate the extent of fractional crystallization required for a particular parent liquid to proceed to that which will precipitate a particular cumulus mineral assemblage the identity of the primary compositions of the cumulus minerals (primocrysts) are needed as the latter may have changed in composition during equilibrium crystallization of trapped liquid.

The Skaergaard layered series cumulus plagioclase crystals consist of homogeneous cores and zoned rims. The cores have been interpreted as primocryst compositions with the rims representing the crystallization product of trapped inter-cumulus liquid[1,7]. However, olivine crystals are virtually homogeneous[8] and must have continuously re-equilibrated, and become less magnesian, during further crystallization of trapped liquid. The extent of the change in primocryst mineral compositions during re-equilibration depends on the proportions of primocrysts and trapped liquid. This may be illustrated by considering chill EG4507 which has olivine(Fo86) and plagioclase(An81) as liquidus phases[9]. In Table 1 compositions B1, B2 and B3 are the result of accumulation of 50% of olivine, 50% of plagioclase and 50% of olivine(15%) + plagioclase(35%) into the original liquid. The ratio of olivine to plagioclase incorporated into EG4507 to form composition B3 is that in which they are predicted to crystallize during fractional crystallization. For these bulk compositions the equilibrium mineral compositions at the solidus can be approximated by the CIPW norm with the anorthite and albite contents recast to indicate plagioclase composition and the Fe–Mg distribution between the ferromagnesian minerals reset using measured cation exchange reaction coefficients[10,11]. Note that the olivine in the bulk composition buffered by the 50% of previously crystallized cumulus olivine is significantly more magnesian than the olivine in the composition buffered by 50% of cumulus plagioclase. Similarly, plagioclase is much more calcic in the bulk composition equilibrated with 50% of cumulus plagioclase. Such variations in mineral composition do occur in layers as narrow as 10 cm (refs 8, 12).

A critical point in the fractional crystallization sequence of the Skaergaard parent liquid is that at which clinopyroxene first appears as a cumulus phase at the base of LZb. Identification of the primocryst mineral compositions at this point could constrain the composition of the coexisting liquid and thus of the range of liquids in the fractionation sequence. From this discussion, and the fact that plagioclases are zoned, the best assumption is that the coexisting primocryst compositions are the most calcic plagioclase and most magnesian ferromagnesian minerals observed at this structural height, not necessarily from the same sample: olivine(Fo64), plagioclase(An63) and clinopyroxene(Di71 = $Mg/(Mg+Fe^{2+})$)[1,6]. The problem is then to identify which of the possible parent liquids would precipitate this mineral assemblage at the point where clinopyroxene joins olivine and plagioclase as fractionating phases.

Theoretical fractional crystallization

To predict magmatic evolution based on known phase equilibria we need to predict liquidus temperatures, liquidus phase compositions and the change in liquid composition produced by perfect fractional crystallization of the liquidus phases. Nathan and Van Kirk[11] presented a generalized method incorporating these features but the accuracy of liquidus temperature equations, relating mineral phase liquidus temperature to liquid composition at atmospheric pressure, have been criticized[17]. Applied to a data set consisting of microprobe analyses of glasses produced in experiments at atmospheric pressure in this laboratory, and for which experiment temperature and oxygen fugacity were known, the equations produce systematic discrepancies of up to 40 °C. The data set has been compiled from experiments on 24 terrestrial basaltic rocks, 5 komatiites and 10 achondritic meteorite and lunar basalts. At present, the data set contains 336 olivine saturated glasses, 296 spinel–magnetite saturated glasses, 160 plagioclase saturated glasses, 116 clinopyroxene saturated glasses and 67 low Ca-pyroxene saturated glasses. With the equation compositional coefficients revised based on this data set, liquidus temperatures for olivine, plagioclase, clinopyroxene and magnetite fit the equations with a standard deviation of error of <6 °C (refs 18, 19). With this degree of accuracy in liquidus temperature prediction it is possible to be much more accurate in the calculation of fractionation paths and more specific in the interpretation of petrogenetic relationships.

The revised coefficients for the orthopyroxene liquidus temperature equation, derived from the orthopyroxene–glass pairs in the data set, give a standard deviation of error of <5 °C (all structural modifications of low-Ca pyroxene are referred to as orthopyroxene). However, the resulting equation predicts that orthopyroxene would be the liquidus phase in glass compositions which did not coexist with orthopyroxene. This anomaly probably results from the limited range of liquid compositions in the data set for orthopyroxene–glass pairs and illustrates the possible consequences of undue extrapolation of the liquidus temperature equations. During fractional crystallization in a multi-component system the extent of extrapolation may not be readily apparent. Until this anomaly is resolved the quantitative application of the mineral temperature equations will be restricted to equilibria not involving orthopyroxene.

The predicted and experimentally determined liquidus temperatures for EG4507 and KT39, neither of which were included in the data set, are presented in Table 2. For KT39 the equations predict that orthopyroxene would be close to the liquidus although in the experiments this phase was not detected after 55 °C of equilibrium crystallization[6]. Apart from this, the rather small discrepancies between predicted and experimentally determined liquidus temperatures may be accounted for by known experimental and analytical problems. The liquidus temperature equations are, therefore, considered sufficiently accurate to demonstrate possible petrogenetic relationships between Skaergaard chill compositions before the appearance of orthopyroxene. It could even be argued that the predictions for KT39 fit the observed petrography[6] better than the experiments.

In many pressure–temperature (*P–T*) phase diagrams of basic rocks clinopyroxene becomes the liquidus phase at elevated pressure. The possible consequences of the involvement of clinopyroxene in fractionation equilibria at high pressure can be considered after assessment of the effects of total pressure on the liquidus temperature equations. The effects of total pressure can be estimated approximately from the experimentally determined slopes of liquidus surfaces on *P–T* phase diagrams of rocks[21–25]. Average values of +4 °C kbar^{-1} for olivine, plagioclase and magnetite and +9 °C kbar^{-1} for clinopyroxene and orthopyroxene have been adopted. Similarly the effects of small quantities of dissolved water can also be estimated from data in the systems albite–water and diopside–anorthite–water[13]. Values of −250 °C × H^+ (cation fraction) for felsic minerals and −80 °C × H^+ for mafic minerals have been adopted.

Using the mineral temperature equations to identify the liquidus phase the liquid path followed during fractional crystallization can be approximated by repeated removal of a small amount of the current liquidus phase using equations for exact fractionation[14]. For olivine, clinopyroxene and orthopyroxene these are of the form

$$\mathrm{Fe} = (\mathrm{Fe}_0/\mathrm{Mg}_0^{K}).\mathrm{Mg}^{K}$$

where the element values are absolute cation quantities in the liquid before (Fe_0, Mg_0) and after (Fe, Mg) each fractionation step and K is the appropriate cation exchange reaction coefficient[10,11]. The latter may also be used to determine the composition of the liquidus phase. The amounts of other components to be removed from the liquid are determined by the stoichiometry of the phase in question. For plagioclase and magnetite the equations for exact fractionation can be solved by numerical approximation:

Plagioclase:

$$\mathrm{Na}^{((\mathrm{K.Ca.Al})/\mathrm{Si})} = (\mathrm{Ca}.\mathrm{Na}_0{}^{((\mathrm{K.Ca_0.Al_0})/\mathrm{Si_0})})/\mathrm{Ca}_0$$

Magnetite:

$$\mathrm{Ti}^{((\mathrm{K.Fe^{3+}})/(\mathrm{Fe^{2+}.Ti}))} = (\mathrm{Fe}^{3+}.\mathrm{Ti}_0{}^{((\mathrm{K.Fe^{3+}}_0)/(\mathrm{Fe^{2+}}_0.\mathrm{Ti}_0))})/\mathrm{Fe}^{3+}{}_0$$

At a phase boundary, crystallization of one phase results in the liquid moving into the primary phase volume of one of the other phases present. This then becomes the liquidus phase for the next fractionation step. A cotectic is thus followed by a zig-zag approximation path which is almost independent of the size of the crystallization increment provided that the increments are sufficiently low. At a reaction boundary the phase in reaction fails to become a liquidus phase and ceases to fractionate. Specification of a negative increment allows a liquid fractionation path to be followed back to higher temperature, more primitive compositions. This process can be described as negative fractional crystallization to avoid conflict of terminology with melting processes.

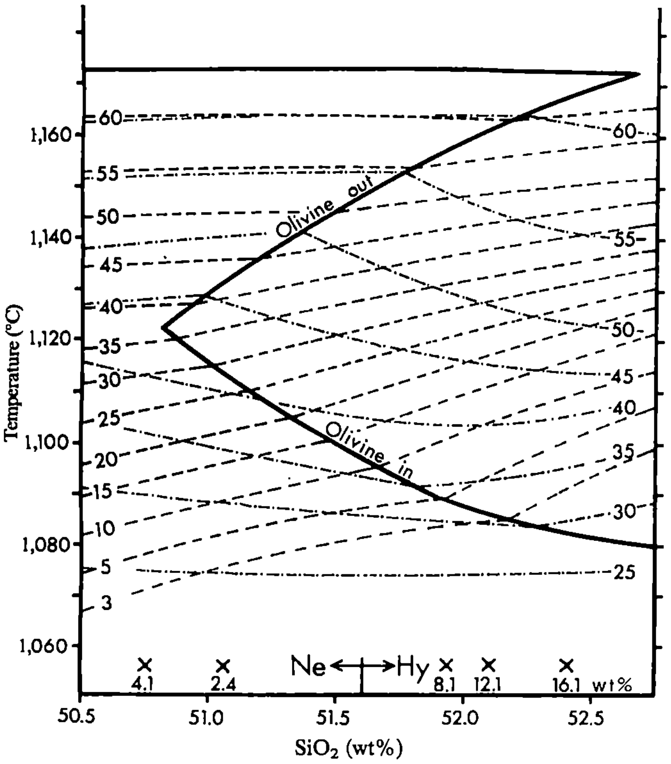


Fig. 2 Fractional crystallization diagram based on the liquidus-temperature equations from ref. 11 for the univariant range of liquids in the system SiO_2–Al_2O_3–FeO–MgO–CaO–Na_2O which all have olivine(Fo64), plagioclase(An63) and clinopyroxene(Di71) as simultaneous liquidus phases. Vertical lines illustrate the changing primocryst compositions and assemblages which would be produced by fractional crystallization. Dashed lines are isopleths of Fo content of olivine, shown with metastable continuation through the field in which olivine is absent, but orthopyroxene is present, while dashed-dotted lines are isopleths of An content of plagioclase. The nepheline and hypersthene scales are not linear.

The liquidus temperature and fractionation equations have been incorporated into a computer program. Given a parent liquid composition the program has been designed to list the compositions and proportions of crystallizing phases and the evolved liquid composition at regular intervals during the fractionation process. The compositions of crystallizing phases correspond to the primocryst mineral compositions of cumulate sequences.

Crystallization of Skaergaard chills

At atmospheric pressure, 37% of fractional crystallization of olivine and plagioclase from EG4507 is predicted to occur before the appearance of clinopyroxene. Only 3% of fractional crystallization of olivine from KT39 is predicted to occur before plagioclase and clinopyroxene appear almost simultaneously. In both cases the ferromagnesian minerals are more magnesian and plagioclases more calcic than the quoted primocryst compositions at the base of LZb. Neither composition could apparently be directly parental to the layered series by fractional crystallization at low pressure unless re-equilibration of primocrysts with trapped liquid has resulted in an incorrect identification of the primocryst compositions.

EG4507 and KT39 are shown in Fig. 1 as an M2S projection into the plane CS-MS-A (ref. 15). The liquid path followed during fractional crystallization of EG4507 at low pressure does not pass through, or even close to, the composition of KT39. Note specifically the orientation of the liquid path followed during fractional crystallization of olivine + plagioclase + clinopyroxene. This is significantly different from the path followed during equilibrium crystallization of these phases. Least squares mixing calculations using the method of Wright and Docherty[16] demonstrate that KT39 could be combined with the olivine, plagioclase and clinopyroxene compositions present close to the liquidus to produce EG4507 (45.1%KT39 + 12.6% olivine(Fo74) + 35.6% plagioclase(An73) + 6.7% clinopyroxene(Di79) = EG4507 sum $r^2 = 0.5$). KT39 could thus be a liquid produced during equilibrium crystallization of a liquid such as EG4507 before separation of crystals from liquid. However, this is not an essential relationship.

Cumulate sequence parent liquids

Although it might be possible to find a chill composition able to produce the observed layered series primocryst assemblages during fractional crystallization it is possible to adopt a radically new approach to the calculation of parent liquid compositions.

Let an olivine–gabbro mineral assemblage consist of primocrysts of olivine (Fo64), plagioclase (An63) and clinopyroxene

Table 2 Comparisons of experimentally determined and calculated liquidus temperatures (°C) for Skaergaard chills at atmospheric pressure

Mineral	EG4507 (fO_2 = iron/wustite)*			KT39 (fO_2 = magnetite/wustite)*		
	XP†	F‡	N & V§	XP‖	F‡	N & V§
Olivine	1,230–1,248	1,225	1,241	1,171–1,179	1,179	1,187
Plagioclase	1,230–1,248	1,219	1,220	1,171–1,174	1,167	1,165
Magnetite	<1,174	1,193	—	1,171–1,179	1,154	1,142
Clinopyroxene	1,188–1,208	1,173	1,213	1,134–1,140	1,157	1,181
Orthopyroxene	<1,174	1,181	1,166	<1,115	1,176	1,170

The liquidus-temperature predictions for phases which would not actually be the liquidus phases cannot be compared with the temperature of appearance of these phases during equilibrium crystallization. They are given for comparison and to give an approximate indication of the possible effect of total pressure (see text).

* Fe_2O_3 and FeO contents recalculated using an equation given in ref. 11.

† From ref. 9. Upper temperature is that at which phase was absent, lower temperature is that at which phase was present.

‡ F, prediction based on data set of experimental glasses.

§ N & V, prediction of the equations given in ref. 11.

‖ From ref. 6, magnetite liquidus may be too high.

(Di71), the assemblage marking the base of LZb. In the system SiO_2–Al_2O_3–FeO–MgO–CaO–Na_2O one liquid which would precipitate this assemblage would lie in the pseudo-ternary plane olivine(FoX)–plagioclase(AnY)–clinopyroxene(DiZ), where X, Y and Z represent the mineral compositions which form the plane end-members. Values for X ($=Z$) can be calculated from crystal–liquid iron–magnesium cation exchange reaction coefficients[10,11] and the identified primocryst ferromagnesian mineral compositions. The value of Y depends on the proportions of plane end-members contributing to the liquid composition at the piercing point where the three crystalline phases are in equilibrium with liquid. It is easy to estimate Y, use the liquidus temperature equations to generate the pseudo-ternary phase diagram, observe the composition of the plagioclase in equilibrium with the ferromagnesian minerals of the selected composition, and refine Y accordingly. The liquid composition is represented by the plane end-members at the piercing point determined to closer than 0.1 wt%.

The resulting liquid composition is not unique. In a six-component system the phase rule is uniquely satisfied by the coexistence of four phases and the specification of pressure and the compositions of three of the phases. However, in this instance the compositions of olivine and both pyroxenes are determined by the Mg/(Mg+Fe) ratio of the liquid. Therefore, this mineral assemblage could be precipitated from a univariant range of liquid compositions in the six-component system. The extent of this range can be determined by altering the concentration of any one of the components of plagioclase in the parent liquid. For convenience SiO_2 has been chosen because the effect of varying the concentration of this component is easily expressed in terms of normative nepheline, olivine, hypersthene or quartz.

Figure 2 is a fractional crystallization diagram in which the specified mineral assemblage is crystallized at the liquidus over a range of nepheline and hypersthene normative compositions. Note that different liquids would produce progressively differing primocryst assemblages as illustrated by the mineral composition isopleths and the location of the boundaries defining the field in which olivine is absent. This is the key to the interpretation of natural igneous cumulate assemblages. The extent of the orthopyroxene field and its extension into nepheline normative compositions are almost certainly erroneous, resulting from extrapolation of the orthopyroxene liquidus temperature equation. Note also that the extent of the 'olivine gap' is a function of the SiO_2 content of the parent liquid.

Adding more components to the system will produce a greater range of liquids capable of precipitating the same mineral assemblage. However, it should be possible to restrict element abundances in parent liquids from consideration of crystal–liquid partition coefficients or from element abundances in extrusive rocks. The one liquid which is best able to reproduce the observed sequence can be obtained by adjustment of the minor components by trial and error. To promote the appearance of cumulus magnetite at the correct point in the fractionation sequence in relation to the compositions of the coexisting primocrysts at the base of LZc there is comparatively little room for adjustment of the contents of K_2O, TiO_2 and Fe_2O_3 in the parent liquid. However, as shown in Fig. 2, it is the point of disappearance and reappearance of olivine, as a result of reaction relationships with orthopyroxene, and the compositions of primocrysts in the later stages of fractional crystallization which identify more clearly the most probable parent liquid composition. Refinement of the liquid composition present at the base of LZb should, therefore, await additional experimental data on orthopyroxene–liquid equilibria.

Negative polybaric fractional crystallization

For any given liquid composition a range of parental compositions can be generated by negative fractional crystallization. For EG4507, for example, parent liquids may lie some way back along the olivine–plagioclase cotectic and then even further back towards olivine alone. Alternatively at high pressure, parent liquids may lie along cotectics involving olivine+plagioclase+clinopyroxene, or olivine+clinopyroxene and ultimately olivine alone.

Two particular features emerge from these general considerations. (1) Provided that the same phases are involved at all stages, negative fractional crystallization of a range of compositions tends to produce convergence of their parental compositions. The small but significant differences in composition which would be amplified by subsequent positive fractional crystallization are more easily produced in converged parental compositions. For example, Ca/Al ratios are very sensitive to crystallization, or otherwise, of clinopyroxene. For this reason some of the range of Skaergaard chill compositions might be more easily related to each other by equilibria at very high pressure. (2) Polybaric fractional crystallization of a particular assemblage of phases from a particular liquid, accomplished by lowering pressure by a small selected amount after each increment of crystallization, may produce residual liquid compositions significantly different from the liquid compositions produced by isobaric fractional crystallization of the same phases. For example, positive isobaric fractional crystallization of olivine+plagioclase+clinopyroxene from EG4507 at 10 kbar produces residual liquids with Al_2O_3 contents rising to over 20% whereas polybaric fractional crystallization of these phases from 10 to 1 kbar produces residual liquids with Al_2O_3 contents falling to below 16%.

Conclusions

The integration of data from phase equilibrium experiments on a range of bulk rock compositions by means of a set of mineral-phase liquidus-temperature equations is sufficiently accurate to allow some quantitative modelling of crystal–liquid fractionation processes. The accuracy of this approach and the scope for extending its range of applicability to higher pressures and to equilibria involving additional phases such as garnet and amphibole are being examined. This approach should be able to test for possible relationships between associated igneous rocks; to determine the conditions of formation of phenocryst assemblages; and to constrain cumulus sequence parental liquid compositions in cases for which no appropriate chilled liquid sample is available. It is also easy to incorporate trace elements into the fractionation scheme.

I thank Drs K. G. Cox and G. M. Biggar for valuable comments and D. G. Russell for his unpublished glass analyses.

Received 31 July 1980; accepted 6 February 1981.

1. Wager, L. R. & Brown, G. M. *Layered Igneous Rocks* (Oliver & Boyd, Edinburgh, 1967).
2. McBirney, A. R. *Nature* **253,** 691–694 (1975).
3. Irvine, T. N. in *The Evolution of the Igneous Rocks, 50th Anniversary Perspectives* (ed. Yoder, H. S. Jr) 245–306 (Princeton University Press, 1979).
5. Blank, H. R. & Gettings, M. E. *Trans. Am. Geophys. Un.* **54,** 507 (1973).
6. Hoover, J. D. *Yb. Carnegie Instn Wash.* **77,** 732–743 (1978).
7. Maaløe, S. *J. Petrol.* **17,** 398–419 (1976).
8. Nwe, Y. Y. *Contr. Miner. Petrol.* **55,** 105–126 (1976).
9. Biggar, G. M. *Contr. Miner. Petrol.* **46,** 159–167 (1974).
10. Roedder, P. L. & Emslie, R. F. *Contr. Miner. Petrol.* **29,** 275–289.
11. Nathan, H. D. & Van Kirk, C. K. *J. Petrol.* **19,** 66–94 (1978).
12. McBirney, A. R. & Noyes, R. M. *J. Petrol.* **20,** 487–554 (1979).
13. Burnham, C. W. in *The Evolution of the Igneous Rocks, 50th Anniversary Perspectives* (ed. Yoder, H. S. Jr) 428–439 (Princeton University Press, 1979).
14. Pearce, T. M. *Nature* **276,** 771–774 (1978).
15. O'Hara, M. J. *Earth. Sci. Rev.* **4,** 69–133 (1968).
16. Wright, T. L. & Doherty, P. C. *Bull. geol. Soc. Am.* **81,** 1995–2008 (1970).
17. Hostetler, C. J. & Drake, M. J. *J. geophys. Res.* **85,** 3789–3796 (1980).
18. Russell, D. G. *Progr. exp. Petrol.* (in the press).
19. Ford, C. E. & Russell, D. G. (in preparation).
20. Ford, C. E. *Progr. exp. Petrol.* (in the press).
21. Bender, J. F. *et al. Earth planet. Sci. Lett.* **41,** 277–302 (1978).
22. Bickle, M. J. *et al. Earth planet. Sci. Lett.* **37,** 97–106 (1977).
23. Thompson, R. N. *Contr. Miner. Petrol.* **45,** 317–341, 1974.
24. Ford, C. E. *et al. Geochim. cosmochim. Acta* Suppl. 3 **1,** 207–229 (1973).
25. Fisk, M. R. *et al. Contr. Miner. Petrol.* **74,** 361–374 (1980).

Chloroplast protein phosphorylation couples plastoquinone redox state to distribution of excitation energy between photosystems

John F. Allen[*†], **John Bennett**[†], **Katherine E. Steinback**[*‡] **& Charles J. Arntzen**[*‡]

[*] USDA-SEA-AR, Department of Botany, University of Illinois, Urbana, Illinois 61801, USA
[†] Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, UK
[‡] MSU/DOE Plant Research Laboratory, Michigan State University, East Lansing, Michigan 48824, USA

In photosynthetic membranes isolated from pea leaves, the redox state of the plastoquinone pool controls both the level of phosphorylation of the chloroplast light-harvesting pigment–protein complex (LHC) and distribution of absorbed excitation energy between the two photosystems. Phosphorylation of LHC polypeptides is proposed as the regulatory mechanism by which photosynthetic systems adapt to changing wavelengths of light.

In chloroplasts, non-cyclic photosynthetic electron transport generates ATP and NADPH by the operation of photosystem I and photosystem II acting in series. The energy that drives this process is derived partly from light absorbed by antenna chlorophylls associated closely with each photosynthetic reaction centre and partly from the more peripheral light-harvesting pigments[1] which are organized into specific chlorophyll–protein complexes[2–4]. The most abundant of the chlorophyll-binding proteins in green plants is a 25,000 molecular weight (MW) polypeptide known as the light-harvesting chlorophyll *a/b* protein (LHCP).

In non-cyclic electron transport, the two photosystems must operate at the same rate. If absorbed radiant energy is distributed between the photosystems in such a way that one photosystem is excited more than the other, the excess excitation energy is dissipated, with a consequent loss of photosynthetic quantum efficiency. However, it is clear from studies with algal cells exposed to red and far-red light that chloroplasts can detect such an imbalance in the excitation of the photosystems and at least partially correct it by increasing the proportion of absorbed excitation energy that is distributed to the rate-limiting photosystem[5–8]. These adaptive changes are referred to as state 1–state 2 transitions. State 1 is the adaptive state which results from over-excitation of photosystem I (as, for example, by far-red light) and is characterized by an increase in the proportion of total absorbed excitation energy that is transferred to photosystem II relative to photosystem I. State 2 results from over-excitation of photosystem II by red light and is characterized by increased energy transfer to photosystem I.

The biochemical events underlying state 1–state 2 transitions are poorly understood. However, a related phenomenon that has been extensively studied at the biochemical level is the ability of metal cations to affect the distribution of absorbed excitation energy in isolated thylakoids[9–11]. In the presence of salts of monovalent and divalent metal cations[12], the light-harvesting chlorophyll–protein complexes that contain LHCP are structurally and functionally associated with photosystem II, whereas in the absence of salts the association between the light-harvesting apparatus and the photosystems is altered and excitation energy transfer to photosystem I increases.

A major factor governing the response of thylakoids to cations *in vitro* is the LHC itself. Experiments using thylakoids have shown that trypsin can selectively remove a surface-exposed segment of the LHCP and that this removal is concomitant with the loss of the ability of cations to affect the distribution of excitation energy[13]. In addition, antibodies raised against isolated, purified LHC were found to react with surface-exposed segments of the membrane-bound complex and to prevent such cation effects[14]. These studies indicate that the regulatory mechanism controlling the distribution of absorbed excitation energy in chloroplasts involves a surface-exposed segment of the LHCP.

It has recently been established that the LHCP is reversibly phosphorylated on a threonyl residue located within this surface-exposed segment[15–18]. The LHCP belongs to a group of thylakoid polypeptides shown to be phosphorylated by a membrane-bound, light-activated protein kinase[19,20] and dephosphorylated by a membrane-bound phosphatase[17]. Bennett *et al.*[21] have demonstrated that the reversible phosphorylation of LHCP changes the light-harvesting properties of isolated thylakoids: phosphorylation increases (and dephosphorylation decreases) the proportion of absorbed excitation energy transferred to photosystem I.

This article shows that the activity of the protein kinase is regulated by the redox state of plastoquinone. We propose a molecular mechanism for state 1–state 2 transitions *in vivo* that

0028–0836/81/190025—05$01.00

is consistent with these data and we argue that the presence of cations is merely a prerequisite for these transitions rather than a decisive regulatory factor in itself.

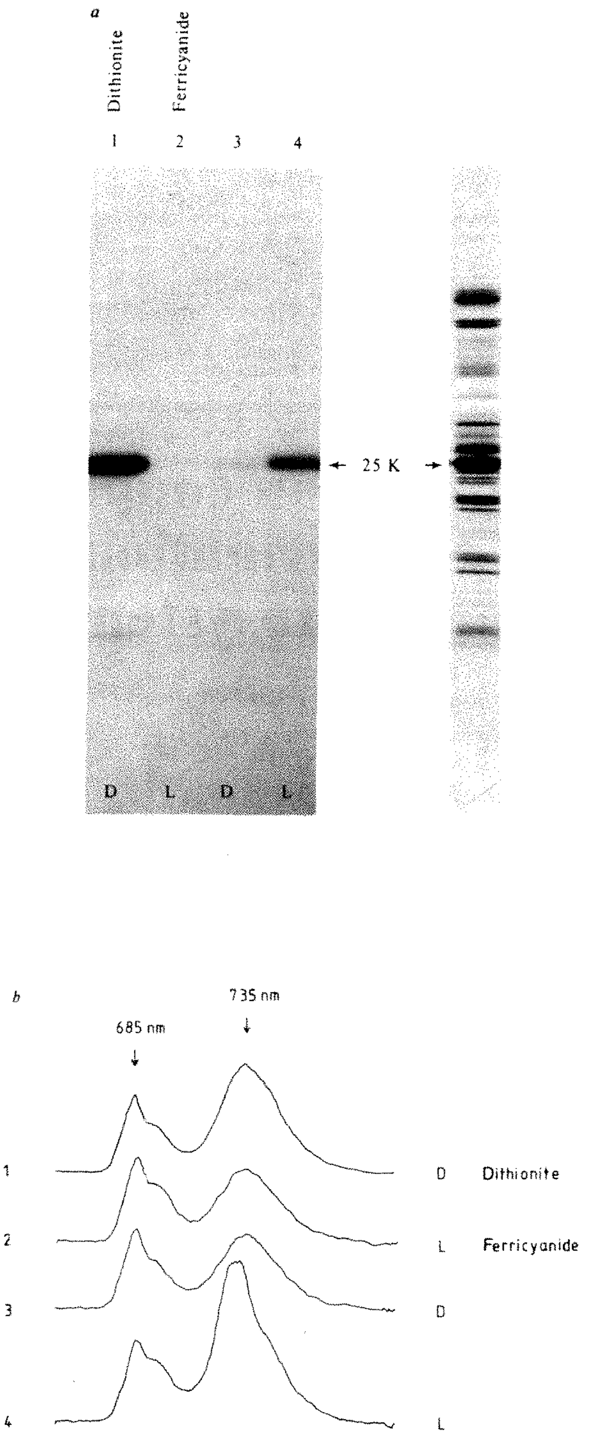


Fig. 1 Effects of light, dithionite and ferricyanide on phosphorylation of LHCP (*a*) and on 77 K fluorescence emission spectrum of thylakoids (*b*). *a*, Thylakoids that had been pre-equilibrated with or without dithionite (10 mM) in darkness and with or without ferricyanide (2 mM) in the light were incubated with [γ-^{32}P]ATP as described in Table 1 legend. Total ^{32}P incorporation into acid-insoluble material was determined and then each sample was subjected to SDS-polyacrylamide gel electrophoresis, followed by staining (lane 5) and autoradiography (lanes 1–4). The identities of the lanes (together with total protein kinase activity of the thylakoids, expressed as a percentage of the light control) are: lane 1, dark + dithionite (124%); lane 2, light + ferricyanide (33%); lane 3, dark (16%); lane 4, light (100%). The position of the 25,000 MW LHCP is indicated (25 K). *b*, Thylakoids were pre-equilibrated and incubated as described above except that the dithionite concentration was 20 mM and the ATP was not radioactive. After incubation, thylakoids were quickly diluted 10-fold with a medium containing 50 mM Tricine-NaOH (*p*H 7.6), 30% (v/v) glycerol and 10 mM $MgCl_2$ and then frozen to 77 K. Fluorescence emission spectra were obtained as described previously[21] using 440 ± 10 nm light for excitation. Spectra were normalized to give equal relative fluorescence at 685 nm using a Hewlett-Packard 9825 computer.

Activation of the thylakoid protein kinase by light and reducing agents

The protein kinase activity of isolated thylakoids was assayed by measuring the transfer of ^{32}P from [γ-^{32}P]ATP to total membrane protein (Table 1). Light stimulated the kinase about fourfold compared with a dark-incubated control. The herbicide 3-(3′,4′-dichlorophenyl)-1,1-dimethylurea (DCMU), which blocks electron transport between the primary acceptor (Q) of photosystem II and plastoquinone, completely abolished the stimulatory effect of light. The stimulation could not be restored by addition of 2,6-dichlorophenolindophenol (DCPIP) in the reduced form ($DCPIPH_2$), which donates electrons to the intersystem chain after plastoquinone. Thus, the kinase seemed to be stimulated by reduction of certain intersystem electron carriers. This idea was supported by the effects of oxidizing and reducing agents on kinase activity (Table 1). Ferricyanide and methyl viologen accept electrons from photosystem I and consequently tend to oxidize the intersystem carriers. These oxidants inhibited light-activation of the kinase. In contrast, the low-potential reductant dithionite[22] activated the kinase in the dark, as did tetramethylhydroquinone ($TMQH_2$, a specific donor to plastoquinone in the intersystem electron transport chain[23,39]). Although dithionite was a somewhat more effective activator than light, their effects were equivalent rather than additive or synergistic. It is likely, therefore, that light and dithionite activated the kinase by the same basic mechanism, that is, by reduction of the intersystem carriers.

When activated by light, the protein kinase of the thylakoid phosphorylates predominantly the 25,000 MW LHCP[16]. Figure 1*a* shows that the same protein was phosphorylated when the kinase was activated in the dark by dithionite. Dithionite also stimulated phosphorylation of the 8,000 MW polypeptide reported to be the proton channel of the thylakoid ATP

Table 1 Activation of thylakoid protein kinase by light and reducing agents

Incubation conditions	Kinase activity (%)
Light	100
Dark	28
Light + DCMU	21
Light + DCMU + $DCPIPH_2$	18
Dark + dithionite	111
Light + dithionite	112
Light + DCMU + dithionite	121
Dark + $TMQH_2$	101
Light + methyl viologen	27
Light + ferricyanide	17

Thylakoid membranes were prepared from 12–20-day-old peas (*Pisum sativum* L. var. Progress no. 9) grown as previously described[3]. Shoots (25 g) were homogenized for 5 s in a blender at low speed in 100 ml of 50 mM Tricine-NaOH (*p*H 7.8), 0.4 M sorbitol and 5 mM $MgCl_2$. The homogenate was filtered through cheesecloth, and centrifuged at 1,000*g* for 5 min. The pelleted chloroplasts were osmotically shocked in 50 ml of 10 mM Tricine-NaOH (*p*H 7.8) and 5 mM $MgCl_2$. The suspension was centrifuged at 2,800*g* for 5 min and the thylakoid pellet was resuspended to a final chlorophyll concentration of 1 mg ml^{-1} in 50 mM Tricine-NaOH (*p*H 7.8), 100 mM sorbitol and 10 mM $MgCl_2$. Portions (100 μl) of the thylakoid suspension were equilibrated at 20 °C in a final volume of 1 ml of the same buffer containing the indicated additions: 3-(3′,4′-dichlorophenyl)-1,1-dimethylurea (DCMU) at 10 μM; 2,6-dichlorophenolindophenol (DCPIP) at 50 μM, maintained in the reduced form ($DCPIPH_2$) by sodium ascorbate (2 mM); sodium dithionite (20 mM); tetramethylhydroquinone ($TMQH_2$, 500 μM) reduced with sodium borohydride as described elsewhere[23]; methyl viologen (50 μM); potassium ferricyanide (2 mM). Equilibration took place either in the light (15 μE m^{-2} s^{-1}) or darkness, and, where dithionite was included, reaction mixtures were made anaerobic by bubbling them with nitrogen gas before the addition of thylakoids. Phosphorylation reactions were started by addition of [γ-^{32}P]ATP (10 Ci mol^{-1}) to a final concentration of 100 μM. The reaction was stopped after 4 min by addition of 0.5 ml of cold trichloroacetic acid (30%, w/v). Acid-stable, precipitable radioactivity was counted by Cerenkov spectrometry. The extent of protein phosphorylation was calculated from the mean activity of duplicated samples. Results are expressed as a percentage of the rate obtained in the light in aerobic conditions (5 nmol phosphate incorporated per mg chlorophyll in 4 min).

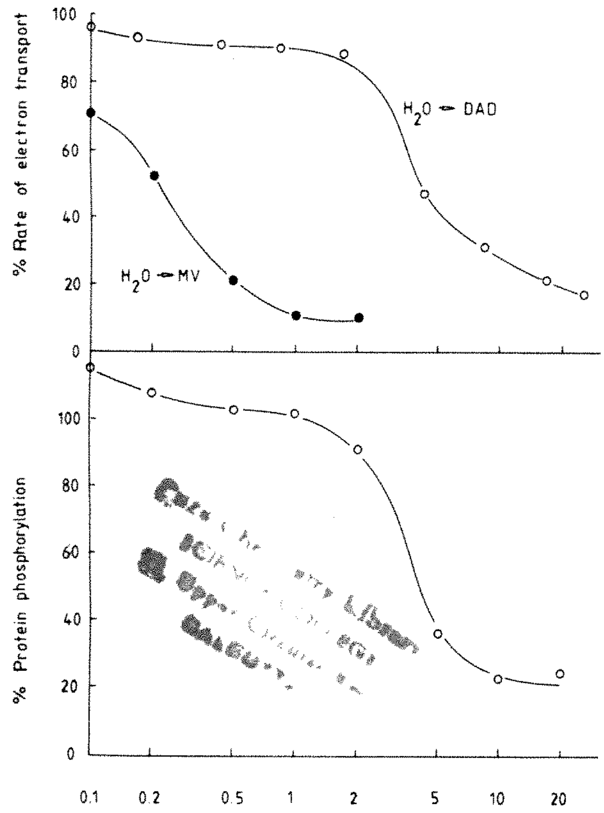


Fig. 2 Inhibition of electron transport and protein kinase activity by 2,5-dibromo-3-methyl-6-isopropyl-*p*-benzoquinone (DBMIB). Whole chain electron transport was measured as oxygen uptake in the presence of methyl viologen (MV, 100 μM), sodium azide (5 mM) and ATP (100 μM). The control rate in the absence of DBMIB was 23 μmol (mg chlorophyll h)$^{-1}$. Photosystem II-mediated electron transport was measured as oxygen evolution in the presence of $K_3Fe(CN)_6$ (2.5 mM), diaminodurene (DAD, 1.2 mM) and ATP (100 μM). The control rate was 97 μmol (mg chlorophyll h)$^{-1}$. All electron transport measurements were carried out at saturating light intensity (1 mE m^{-2}s^{-1}) in a reaction medium (2 ml) containing 50 mM Tricine (*p*H 7.8), 100 mM sorbitol, 10 mM $MgCl_2$ and chloroplasts equivalent to 200 μg chlorophyll. Protein kinase activity was measured as described in Table 1 legend. DBMIB was present during the equilibration period before the addition of [γ-^{32}P]ATP. Values for phosphorylation were calculated by subtraction of the activity of a dark control and are expressed as a percentage of the activity in the light in the absence of DBMIB.

synthase[20]. In contrast, ferricyanide completely inhibited phosphorylation of these thylakoid polypeptides.

Changes in the distribution of absorbed excitation energy

Bennett *et al.*[21] reported that light-induced protein phosphorylation altered the 77 K chlorophyll fluorescence emission spectrum of isolated thylakoids. The finding that emission at 735 nm (F_{735}; derived mainly from the chlorophylls of photosystem I) was increased relative to emission at 685 nm (F_{685}; derived mainly from light-harvesting chlorophylls associated with photosystem II[1]) was one of several lines of evidence[21] indicating that protein phosphorylation changed the distribution of excitation energy within the pigment bed of the membrane in such a way as to favour distribution to photosystem I at the expense of photosystem II.

If the above conclusion is correct, changes in the distribution of absorbed excitation energy should occur whenever the level of protein phosphorylation is altered regardless of how the latter is brought about. Figure 1*b* shows that the 77 K chlorophyll fluorescence emission spectrum of isolated thylakoids was changed by ferricyanide and dithionite in accordance with the effects of these reagents on protein phosphorylation. Ferricyanide inhibited the light-dependent increase in F_{735} while dithionite increased F_{735} in darkness, thus providing evidence for the control of excitation energy distribution by the redox state of the electron transport chain.

Although the dithionite-induced fluorescence change occurred in the presence of ATP, the data in Fig. 1 do not in themselves indicate that the change required ATP or involved protein phosphorylation. To explore this point in more detail, thylakoids were incubated in various conditions in the presence and absence of ATP and dithionite. Values for the ratio F_{735}/F_{685} were calculated from the resulting 77 K fluorescence emission spectra (Table 2). An ATP-dependent increase in the ratio was observed only in conditions which also permitted protein phosphorylation. Thus, there was no significant effect of ATP on fluorescence in the dark or in the light in the presence of DCMU, unless dithionite was also present. It is also clear from Table 2 that dithionite exerted two effects on the F_{735}/F_{685} ratio, one depending on the presence of ATP and the other not. The effect which required both ATP and dithionite represented the change in the distribution of excitation energy that could be attributed to dithionite-activated protein phosphorylation. The effect of dithionite in the absence of ATP may result from a direct alteration in fluorescence yield.

Activation of protein kinase by reduced plastoquinone

The above data raised the possibility that plastoquinone was the electron transport component whose reduction led to kinase activation. This possibility was tested by the use of 2,5-dibromo-3-methyl-6-isopropyl-*p*-benzoquinone (DBMIB), which inhibits electron transport at two sites—it inhibits oxidation of plastohydroquinone at a high-affinity site[24] and reduction of plastoquinone at a secondary, low-affinity site[25]. Figure 2 illustrates both inhibitory activities. At a concentration of 1 μM,

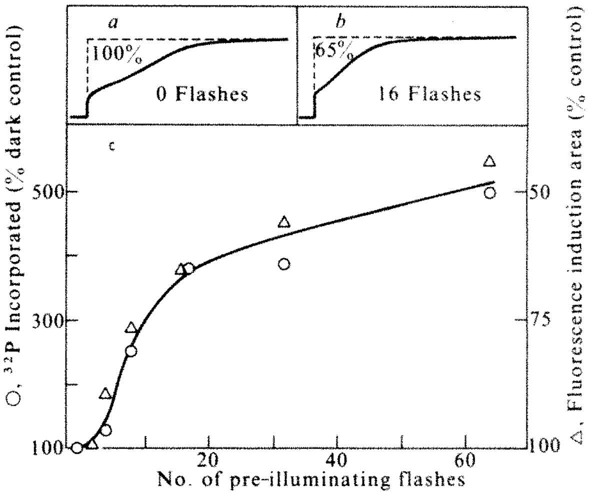


Fig. 3 Thylakoid protein phosphorylation and redox state of plastoquinone as functions of the number of pre-illuminating single-turnover flashes. Thylakoids were diluted with a medium containing 50 mM Tricine (*p*H 7.8), 100 mM sorbitol, 10 mM $MgCl_2$ and 10 mM NaF to a chlorophyll concentration of 50 μg ml^{-1} for protein kinase assays or 5 μg ml^{-1} for room temperature chlorophyll fluorescence induction transients. The dilution medium had been bubbled with nitrogen before the addition of thylakoids. The samples were pre-illuminated with a flashlamp (General Radio 'Stroboslav', type 1539A) delivering one intense, 8-μs flash per s. A fibre-optics bundle carried light from the flash-lamp directly into either cuvettes (for the induction transients) or polyethylene microcentrifuge tubes (for kinase assays). After the indicated number of flashes, fluorescence induction was measured in the cuvettes as previously described[3]. *a,b*, Direct tracings of oscillographic recordings; broken lines delimit the area above the induction curve, a measure of the proportion of the plastoquinone pool that is in the oxidized state. *c*, [γ-^{32}P]ATP was added to the microcentrifuge tubes for a 2-min, dark, anaerobic incubation which was terminated by addition of trichloroacetic acid to a concentration of 10%. Protein phosphorylation was then measured as described in Table 1 legend. Fluorescence measurements and additions of [γ-^{32}P]ATP were made within 5 s of the final pre-illuminating flash.

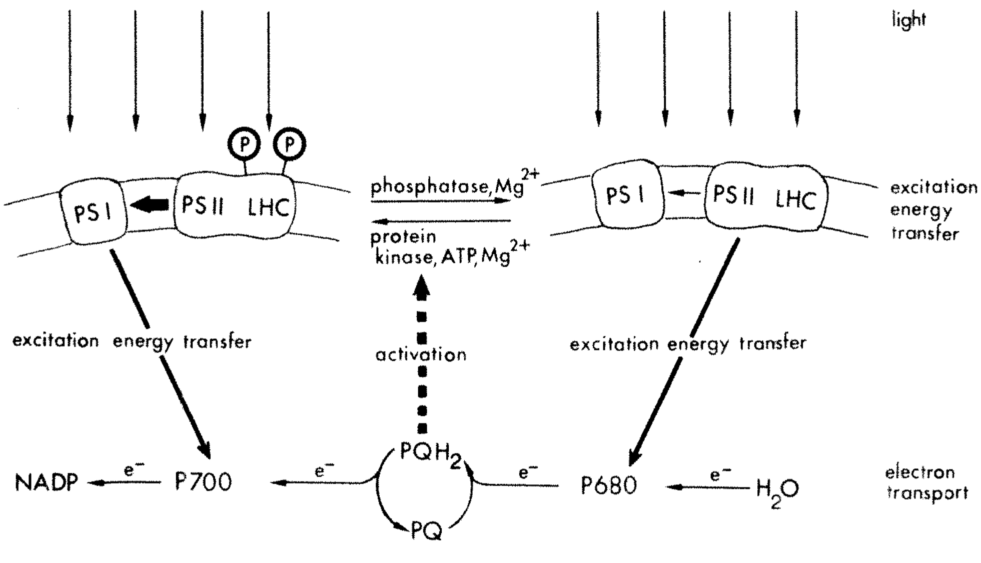


Fig. 4 A model for control of distribution of excitation energy in photosynthesis, in which reversible phosphorylation of LHC couples the redox state of plastoquinone (PQ) to the distribution of excitation energy between the photosystems. The mechanism we propose ensures maximal quantum efficiency of non-cyclic electron transport from water to NADP, mediated by P680 and P700, the reaction centres of photosystem II (PSII) and photosystem I (PSI), respectively. Reduction of plastoquinone by photosystem II leads to kinase activation and LHC phosphorylation, and hence to increased excitation of photosystem I. Conversely, oxidation of plastohydroquinone (PQH_2) by photosystem I inactivates the kinase; the phosphatase then dephosphorylates LHC and hence excitation of photosystem II is increased relative to that of photosystem I. Since the redox state of plastoquinone will be determined in part by the distribution of excitation energy between the two photosystems, reversible LHC phosphorylation completes a feedback loop by means of which any imbalance in this distribution will tend to be self-correcting.

DBMIB inhibited whole chain electron transport (from water to methyl viologen) by 90%, whereas photosystem II-mediated reduction of diaminodurene (DAD) was strongly inhibited only at concentrations greater than about 10 µM (DAD accepts electrons at or before plastoquinone[26]). When light-dependent kinase activity was measured as a function of DBMIB concentration (Fig. 2), strong inhibition was obtained only at concentrations equal to or greater than 10 µM. Thus, the electron transport component that links the activity of the protein kinase to the redox state of the intersystem carriers is situated between the primary and secondary sites of action of DBMIB; the most obvious candidate is plastoquinone itself.

To characterize further the possible relationship between the redox state of plastoquinone and kinase activation, we subjected a series of thylakoid samples to short, intense flashes of light. The redox state of the plastoquinone pool was assessed by recording chlorophyll fluorescence induction transients (Fig. 3). The area above the curve of the transient is proportional to the fraction of the plastoquinone pool that is in the oxidized state[27]. As shown in Fig. 3*a* and *b*, pre-illuminating flashes decreased the area above the curve of the transient, indicating that the plastoquinone pool had accumulated reducing equivalents. In parallel measurements, [γ-^{32}P]ATP was injected in the dark into preflashed samples which were then incubated for 4 min to permit kinase activity assay. As shown in Fig 3*c*, there was a striking similarity between the extent of kinase activation and that of reduction of the plastoquinone pool, as measured by the decrease in the area above the curve of the transient. Thus, the activity of the kinase is linked to the redox state of plastoquinone.

The previous observation[19] that NADPH and ferredoxin could together activate the protein kinase in darkness may now be explained as an indirect effect, resulting from the ability of reduced ferredoxin to reduce plastoquinone[28,29]. We do not believe that activation of the thylakoid-bound protein kinase is mediated by either the thioredoxin system[30] or the 'LEM' system[31], both of which have been implicated in the control of stromal enzyme activity and require electron transport via components on the reducing side of photosystem I.

Table 2 Effects of light, dithionite, DCMU and ATP on the ratio of 77 K fluorescence emissions at 735 nm and 685 nm

Incubation conditions*	F_{735}/F_{685} +ATP (*a*)	−ATP (*b*)	(*a*)/(*b*)†
Dark	1.00	0.95	1.05
Light	1.90	1.20	1.58
Dark + dithionite	1.45	1.10	1.32
Light + dithionite	1.85	1.30	1.42
Light + DCMU	0.85	0.90	0.94
Light + DCMU + dithionite	1.90	1.30	1.46

* As described in Fig. 1*b*; DCMU concentration, 10 µM.

† Values greater than 1 indicate an ATP-dependent change in excitation energy distribution in favour of photosystem I.

Regulation of photosynthetic efficiency by reversible protein phosphorylation

The involvement of plastoquinone in the activation of the thylakoid protein kinase is probably of fundamental importance in the regulation of photosynthesis, for it provides a physiological mechanism for an adaptation which will correct any imbalance in the distribution of excitation energy between the two photosystems (Fig. 4).

The operation of this mechanism may be illustrated by reference to state 1–state 2 transitions. In green algae and higher plants, that is, in photosynthetic organisms containing LHCP, exposure to red light (of wavelength <680 nm) will excite photosystem II to a greater extent than photosystem I, and thereby lead to reduction of plastoquinone. The protein kinase will then be activated, LHCP will become phosphorylated and an increased proportion of absorbed excitation energy will be transferred from the light-harvesting complex to photosystem I at the expense of photosystem II. Thus, the biochemical event underlying the transition to state 2 is phosphorylation of the LHCP. Conversely, under far-red light (of wavelength >680 nm), photosystem I will be excited to a greater extent than photosystem II, plastoquinone will be oxidized and the kinase will be inactivated. Any phosphorylated LHCP will be dephosphorylated by the thylakoid phosphatase and excitation energy transfer to photosystem I will be decreased. Thus, the biochemical event underlying the transition to state 1 is dephosphorylation of the LHCP.

It has often been suggested that state 1–state 2 transitions reflect changes in the cation content of the stroma surrounding the thylakoids[9–11,32], a suggestion based on the effect of cations on the distribution of excitation energy in isolated thylakoids[9–11]. One problem with this idea is that state 2 transitions would have to be correlated with a massive loss of cations from the stroma (to either the thylakoid or the cytoplasm), but there is no reason to believe that the cation concentration of the stroma changes greatly when higher plants or algae are transferred from far-red light to red light or vice versa. However, it should not be assumed that cations are not involved in the regulation of excitation energy distribution *in vivo*. Two roles need to be emphasized. First, cations such as Mg^{2+} ions are required to establish the pattern of excitation energy distribution seen in dephosphorylated membranes, including the close structural and functional association between the LHCP complex and photosystem II[10,12,33]. The second role for Mg^{2+} ions derives from the fact that these cations are cofactors of both the kinase and the phosphatase (Fig. 4)[17,19,20]. Thus, the presence of cations should be thought of as a prerequisite for state 1–state 2 trasitions rather than as a regulatory factor *per se*.

The mechanism depicted in Fig. 4 provides an explanation for the original results of Bonaventura and Myers[5] and of Murata[6] and also for more recent contributions to the study of state 1–state 2 transitions. Several authors have drawn attention to the involvement of the high energy state of the membrane in these transitions[9,34,35]. This would certainly be required for the generation of ATP consumed in protein phosphorylation (transition to state 2). Duysens[36] suggested that the redox state of a component of the intersystem electron transport chain controls these transitions. Plastoquinone would clearly fulfil this role. Horton and Black[37,38] have recently established state 1–state 2 transitions *in vitro*. The transition to state 2 required ATP and was controlled by a membrane-bound factor with the redox potential of plastoquinone. The idea that protein phosphorulation might provide the connection between plastoquinone and ATP is fully in agreement with our present results.

It is important to recognize that state 1–state 2 transitions are also observed in organisms that lack LHCP. For example, Murata[6] and Ried and Reinhardt[8] have shown that these transitions occur in red algae, which contain phycobiliproteins instead of LHCP complexes. Ried and Reinhardt[8] conclude from their studies that the transitions are controlled by the redox state of a component of the intersystem electron transport chain. It will be interesting to discover whether state 1–state 2 transitions in the red algae are brought about by a mechanism analogous to that shown in Fig. 4; that is, a mechanism in which reversible phosphorylation of the phycobiliproteins is controlled by the redox state of plastoquinone.

This study was supported in part by NSF grant PCM 77-18953 to C.J.A. and by SRC grant GR/B11660 to J.B. DBMIB was synthesized and provided by Dr Don Ort. We thank Mrs J. Watson for technical assistance.

Note added in proof: Another piece of evidence for the mechanism in Fig. 4 has now been obtained. Potentiometric redox titrations of plastoquinone, of LHCP phosphorylation and of ATP-dependent chlorophyll fluorescence quenching give identical results[40].

Received 22 December 1980; accepted 18 February 1981.

1. Butler, W. *A. Rev. Pl. Physiol.* **29,** 345–378 (1978).
2. Thornber, J. P., Markwell, J. P. & Reinman, S. *Photochem. Photobiol.* **29,** 1205–1216 (1979).
3. Burke, J. J., Ditto, C. L. & Arntzen, C. J. *Archs Biochem. Biophys.* **187,** 252–263 (1978).
4. Mullet, J. & Arntzen, C. J. *Biochim. biophys. Acta* **589,** 100–117 (1980).
5. Bonaventura, C. & Myers, J. *Biochim. biophys. Acta* **189,** 366–386 (1969).
6. Murata, N. *Biochim. biophys. Acta* **172,** 242–251 (1969).
7. Williams, W. P. in *Primary Processes in Photosynthesis* (ed. Barber, J.) 99–147 (Elsevier, Amsterdam, 1977).
8. Ried, A. & Reinhardt, B. *Biochim. biophys. Acta* **592,** 76–86 (1980).
9. Bennoun, P. & Jupin, H. *Proc. 3rd Int. Congr. Photosynthesis* **1,** 163–169 (1975).
10. Barber, J. *FEBS Lett.* **118,** 1–10 (1980).
11. Vernotte, C., Briantais, J. M., Armond, P. & Arntzen, C. J. *Pl. Sci. Lett.* **4,** 115–123 (1975).
12. Arntzen, C. J. & Ditto, C. L. *Biochim. biophys. Acta* **449,** 259–274 (1976).
13. Steinback, K. E., Burke, J. J. & Arntzen, C. J. *Archs Biochem. Biophys.* **195,** 546–557 (1979).
14. Steinback, K. E., Burke, J. J., Mullet, J. E. & Arntzen, C. J. in *Chloroplast Development* (eds Akoyunoglou, G. & Argyroudi-Akoyunoglou, J. H.) 389–400 (Elsevier, Amsterdam, 1978).
15. Bennett, J. *Nature* **269,** 344–346 (1977).
16. Bennett, J. *Eur. J. Biochem.* **99,** 133–137 (1979).
17. Bennett, J. *Eur. J. Biochem.* **104,** 85–89 (1980).
18. Beliveau, R. & Bellemare, G. *Biochem. biophys. Res. Commun.* **88,** 797–803 (1979).
19. Bennett, J. *FEBS Lett.* **103,** 342–344 (1979).
20. Alfonzo, R., Nelson, N. & Racker, E. *Pl. Physiol.* **65,** 730–734 (1980).
21. Bennett, J., Steinback, K. E. & Arntzen, C. J. *Proc. natn. Acad. Sci. U.S.A.* **77,** 5253–5257 (1980).
22. Ikegami, I. & Katoh, S. *Archs Biochem. Biophys.* **150,** 9–14 (1972).
23. Izawa, S. & Pan, R. L. *Biochem. biophys. Res. Commun.* **83,** 1171–1174 (1978).
24. Hauska, G. in *Encyclopedia of Plant Physiology* Vol. 5 (eds Trebst, A. & Avron, M.) 253–267 (Springer, Berlin, 1977).
25. Gould, J. M. & Izawa, S. *Eur. J. Biochem.* **37,** 185–192 (1973).
26. Izawa, S. in *Encyclopedia of Plant Physiology*, Vol. 5 (eds Trebst, A. & Avron, M.) 268–282 (Springer, Berlin, 1977).
27. Murata, N., Nishimura, M. & Takamiya, A. *Biochim. biophys. Acta* **120,** 23–33 (1966).
28. Arnon, D. I. & Chain, R. K. *Proc. natn. Acad. Sci. U.S.A.* **72,** 4961–4965 (1975).
29. Mills, J. D., Mitchell, P. & Barber, J. *Photobiochem. Photobiophys.* **1,** 3–9 (1979).
30. Wolosiuk, R. A. & Buchanan, B. B. *Nature* **266,** 565–567 (1977).
31. Anderson, L. E. & Avron, M. *Pl. Physiol.* **57,** 209–213 (1976).
32. Murata, N. *Biochim. biophys. Acta* **189,** 171–181 (1969).
33. Butler, W. L. *Brookhaven Symp. Biol.* **28,** 338–346 (1976).
34. Murata, N. & Sugahara, K. *Biochim. biophys. Acta* **189,** 182–192 (1969).
35. Krause, G. H. *Biochim. biophys. Acta* **292,** 715–728 (1973).
36. Duysens, L. N. M. *Biophys. J.* **12,** 858–863 (1972).
37. Horton, P. & Black, M. T. *FEBS Lett.* **119,** 141–144 (1980).
38. Horton, P. & Black, M. T. *Biochim. biophys. Acta* **635,** 53–62 (1981).
39. White, C. C., Chain, R. K. & Malkin, R. *Biochim. biophys. Acta* **502,** 127–137 (1978).
40. Horton, P., Allen, J. F., Black, M. T. & Bennett, J. *FEBS Lett.* (in the press).

IgG antibodies to phosphorylcholine exhibit more diversity than their IgM counterparts

Patricia J. Gearhart
Department of Embryology, Carnegie Institution of Washington, 115 W. University Parkway, Baltimore, Maryland 21210, USA
Nelson D. Johnson, Richard Douglas & Leroy Hood
Division of Biology, California Institute of Technology, Pasadena, California 91125, USA

An amino acid sequence analysis of the N-terminal immunoglobulin heavy and light chain variable regions (V_H and V_L) from 16 hybridoma proteins which bind phosphorylcholine as well as the complete sequence analysis of 9 of these V_H regions is presented. There seem to be more V_H regions participating in the phosphorylcholine response than can be encoded directly by germ-line V_H gene segments. Moreover, the V regions from IgG antibodies are considerably more variable than those from their IgM counterparts. These observations raise the possibility that a somatic mechanism for V region diversification produces greater diversity in IgG than in IgM antibodies.

ONE goal in molecular immunology is to understand the mechanisms responsible for antibody gene diversification. Insight into antibody diversity has been obtained by comparing sequence information from both immunoglobulin genes and proteins[1–8], and our present level of understanding may be summarized as follows. (1) Antibody molecules are composed of heavy (H) and light (L) chains, each of which contain variable (V) and constant (C) regions. (2) Comparative analyses of many variable regions have demonstrated that amino acid sequence variability is unevenly distributed throughout the V region. The antigen-binding site is composed of extremely variable sections termed hypervariable regions (three for each chain), whereas

0028-0836/81/190029—06$01.00

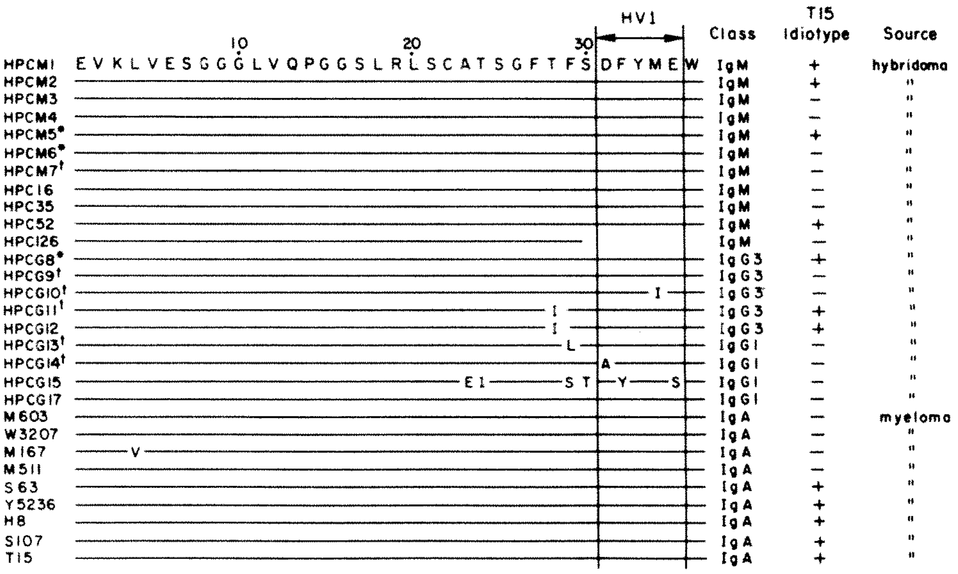


Fig. 1 V_L regions of anti-phosphorylcholine antibodies. The hybridomas were made by the following procedure. BALB/cJ mice (Jackson Laboratory) were immunized intraperitoneally with 0.1 mg of phosphorylcholine–haemocyanin[53] in complete Freund's adjuvant, and 1–2 months later (4 days before cell fusion), injected intraperitoneally with 0.1 mg of phosphorylcholine–haemocyanin in complete Freund's adjuvant. The cells from individual spleens were fused to non-secreting SP2/0-Ag14 cells[54] according to Kennett *et al.*[55]. The heavy chain class of anti-phosphorylcholine antibodies was determined by radioimmunoassay[32], and antibody with the T15 idiotype was detected by binding to A/J anti-T15 serum as described[53]. The HPCM7(1F8), HPCG10(1B6), HPCG13(2F5), HPCG9(2B4), HPCG14(2C5), HPCG15(1B10), HPCG17(24B5) and HPCG12(1D8) hybridomas were obtained from J. Hurwitz[56]. The HPCM3(63b4) and HPCM4(64a5) hybridomas were obtained from B. Clevinger and J. Davie, and were generated from adult mice which were suppressed for expression of the T15 idiotype by treatment with anti-T15 serum at birth. The numbers in parentheses refer to the nomenclature of the donor laboratory. The antibodies were purified and chain separated as previously described[57,58]. The amino acid sequence analysis was carried out on a modified Beckman 890B sequenator as previously described[59]. Amino acids are noted by the one-letter code of Dayhoff[60]. HV1 designates the first hypervariable region which ends at residue 40 (ref. 35). All * proteins come from one mouse, and the + proteins from a second mouse. Other proteins each come from separate individuals. The V_L sequences are categorized according to the T15, M603 and M167 groups with the most typical protein of each group used as the prototype. A solid line indicates identity with the prototype sequence. The sequence of HPC16 is written assuming Asx residues are identical to prototype asparagine or aspartic acid residues. The sequences of HPC52, HPC126, HPC19, HPC104 and HPC16 are from ref. 61; H8, S107, T15 and M603 from ref. 22; S63, W3207 and M511 from ref. 62; and M167 is from ref. 63. Y5236 sequence is from J. Hubert (personal communication).

non-binding portions, termed the framework regions, are more highly conserved[9–12]. The hypervariable regions fold to constitute the walls of the antigen-binding site. (3) The V regions fall into sets of closely related sequences termed groups (formerly called subgroups)[13–15]. The minor variation within the V regions of a particular group has afforded special insight into mechanisms of diversity[2,3,5–8,16,17]. (4) The antibody molecules are encoded by three unlinked families of genes—two light chain families, lambda (λ) and kappa (κ), and a heavy chain family. (5) The V region of the light chain (V_L) is encoded by two distinct gene segments—V_L and J_L (joining)[18], and the V region of the heavy chain (V_H) is encoded by three distinct gene segments—V_H, D (diversity) and J_H (ref. 8). The J_L gene segments encode only a small portion of the third hypervariable region of the light chain, whereas the D and J_H gene segments encode most of the third hypervariable region of the heavy chain. (6) There are at least 100 germ-line V_L and V_H gene segments[17,19], indicating that germ-line diversity is extensive. (7) There are four functional J_κ and J_H gene segments[6–8,20], and it seems that any J_κ gene segment may be joined with any V_κ gene segment[5]. Likewise, the combinatorial joining of V_H, D and J_H gene segments is an important mechanism for amplifying diversity in the third hypervariable region of the heavy chain[3]. (8) The site-specific recombination mechanism which joins the V_L and J_L, or the V_H, D and J_H gene segments, may join particular gene segments at different points in their sequences, thus generating hybrid codons as well as codon insertions or deletions. Thus, alternative joining points generate somatic variation at these junctions during the act of gene segment joining[4,6–8]. (9) The V_L regions seem to exhibit somatic variation in hypervariable and framework regions beyond that which can be explained by the germ-line V gene repertoire and the junctional diversity of V–J joining. In the V_λ region, this diversity appears exclusively in the hypervariable regions[1], but in V_κ regions, diversity is seen in framework regions as well[5].

Table 1 Affinity of antibodies for phosphorylcholine

Antibodies	Class	T15 idiotype	$K_a \times 10^5\ M^{-1}$
T15-V_H, T15-V_L groups			
HPCM2	IgM	+	3.5
HPCM5	IgM	+	2.4
HPCM7	IgM	−	3.2
HPCG8	IgG3	+	1.3
HPCG14	IgG1	−	4.4
T15	IgA	+	4.3
T15-V_H, M603-V_L groups			
HPCM3	IgM	−	0.34
HPCG15	IgG1	−	0.13
W3207	IgA	−	2.5
T15-V_H, M167-V_L groups			
HPCG10	IgG3	−	0.25
HPCG13	IgG1	−	1.8
HPCG17	IgG1	−	1.0
M167	IgA	−	1.9

The affinity of antibodies for phosphoryl[*Me*-^{14}C]choline (NEN) was measured by equilibrium dialysis in a buffer containing 0.17 M boric acid, 0.03 M sodium borate, 0.15 M sodium chloride, *p*H 8.0. Equilibrium dialysis was performed at 25 °C in 0.1-ml compartments at an antibody concentration of 1.5–3.5 mg ml^{-1}. Protein molarity was determined by absorbance at 280 nm using $E^{0.1\%}_{1\,\mathrm{cm}} = 1.37$, and a molecular weight of 150,000 for IgG and IgA, and 170,000 for IgM. The data were plotted by the method of Scatchard[67] to calculate number of sites and by the Sips distribution function[68] to calculate K_a. A least-squares fit was used to obtain the best line. The Scatchard plots were linear, and the number of binding sites ranged from 1.7 to 2.0 for monomeric forms. Values of α, the heterogeneity index, were calculated by the Sips distribution function and ranged from 0.9 to 1.06, indicating homogeneity of binding sites. Binding constants were accurate to ± 10%.

To assess the patterns of variable region diversity to a simple hapten, we have used the hybridoma technique of Köhler and Milstein[21] to obtain homogeneous antibodies from BALB/c mice immunized against phosphorylcholine. Antibodies binding to phosphorylcholine were studied for the following reasons. First, the variable regions of a closely related set of myeloma proteins and serum antibodies that bind phosphorylcholine have been sequenced[15,22–24], and these data can be compared with the sequences obtained from hybridoma antibodies. Second, anti-idiotypic sera raised against several of the myeloma proteins have revealed that the majority of anti-phosphorylcholine antibodies from BALB/c mice have the variable region idiotype of the myeloma protein T15 (refs 25, 26); this is denoted the T15 idiotype. A variety of interesting analyses have subsequently been carried out on the developmental expression[27], genetic mapping[28,29] and regulation[30–32] of this predominant idiotype.

We report here the N-terminal amino acid sequences for the V_H and V_L regions of 16 hybridoma antibodies which bind phosphorylcholine. In addition, we have determined the complete V_H region sequences for nine of these hybridoma antibodies. The most striking observation is that the IgG antibodies exhibit far more diversity than their IgM counterparts. The data also suggest that somatic variation may be involved in V region diversification.

Hybridoma and myeloma immunoglobulins belong to the same V_L and V_H groups

The N-terminal amino acid sequences of 30 V_L regions from hybridoma and myeloma proteins binding phosphorylcholine are given in Fig. 1. Immunoglobulin V_L regions which differ by two or less amino acids over the first 23 residues are classified as members of a single group[15]. The data indicate that the 30 V_L regions fall into three groups. These groups are the same as those initially delineated by analysis of light chains from myeloma proteins and are known as the T15, M603 and M167 groups[33]. Light chains derived from antibodies to phosphorylcholine generated in mouse strains other than BALB/c also fall into these same three V_L groups[34]. Each V_L group has a prototype sequence that is repeated identically many times (Fig. 1). Within a group, each V sequence that differs from the prototype will be denoted a nonprototype sequence.

Of 15 V_L regions in the T15 group, 14 are identical and the 15th differs by a single residue substitution which can be explained by a single nucleotide change. Three of six V_L regions in the M603 group are identical and the other three differ by two, two and five residues, respectively. All these substitutions can be accounted for by single nucleotide changes. Four of nine V_L regions in the M167 group are identical and the others differ by a single residue. Three of these substitutions require two nucleotide changes (HPCG10, HPCG9 and HPC16). Seven of the 15 substitutions noted in these groups fall outside the first hypervariable region as defined by Kabat *et al.*[35]. There seems to be more diversity in the M603 and M167 V_L groups than in the T15 group.

The N-terminal amino acid sequences of the V_H regions from 11 IgM hybridomas, 9 IgG hybridomas and 9 IgA myeloma proteins that bind phosphorylcholine are shown in Fig. 2. Immunoglobulin V_H regions which differ by two or less amino acids over their N-terminal 27 residues belong to a single V group[15]. As seen in Fig. 2, all 29 V_H regions fall within a single V_H group. Twenty-two of the V_H sequences are identical, and six of the remaining seven V_H regions differ by a single residue from the prototype T15 sequence. The remaining V_H region, HPCG15, differs by six residues. Eleven of 12 substitutions can be explained by single nucleotide substitutions, and 8 of 12 substitutions lie outside the first hypervariable region.

The complete V_H region sequences from 10 hybridoma antibodies and 9 myeloma proteins are compared in Fig. 3. The hybridoma and myeloma V_H regions exhibit similar diversity patterns by several criteria. (1) Three hybridoma and four myeloma V_H regions are identical to the T15 prototype V_H sequence. (2) The variant hybridoma V_H sequences differ by 1–6 substitutions and the variant myelomas differ by 1–13 substitutions. (3) In both the hybridoma and myeloma V_H regions, insertions and deletions occur only in the third hypervariable region. (4) In the hybridoma V_H regions, 20 of 25 (80%) variant residues are found in the three hypervariable regions, and similarly, in the myeloma V_H regions, 21 of 28 (75%) variant residues are located in the three hypervariable regions. Since the hybridoma and myeloma V_H regions exhibit the same basic patterns of variation, we will consider both sets of data together in analysing diversity in the V_H segments (residues 1–101), the D segments (residues 102–108) and the J_H segments (109–125), which are encoded respectively, by the V_H, D and J_H gene segments[8].

The V_H segments demonstrate diversity in both framework and hypervariable regions

Ten V_H segments are identical and nine others differ by one to eight substitutions (Fig. 3). Generally, the variant V segments differ by two or more widely scattered substitutions. Note that each variant V segment is distinct from all the others. Twenty-four variant residues are present at 16 different positions. Four positions show variation in more than one V_H region: position 56 has four different amino acid substitutions, position 55 has two different substitutions, position 28 has two identical substitutions and position 40 has three substitutions, two of which are identical. Twelve of 24 substitutions (50%) fall outside the two hypervariable regions contained in the V segment. Single base changes can account for 22 of 24 amino acid substitutions, and two substitutions require two base changes. Thus, any diversification mechanisms must account for diversity outside as well as within hypervariable regions[22].

D segment diversity is extensive

Figure 3 demonstrates that the D segment diversity is extensive and falls into three categories: isolated single base changes (for example, position 106 in M167 and 107 in H8 and HPCG8), substitutions, insertions or deletions at either end of the D segment (as in HPCG13, HPCG14, M603, M511 and M167), and blocks of amino acid substitutions (for example, W3207, HPCM4 and HPCM6). The single base substitutions could arise

Fig. 2 The V_H regions of anti-phosphorylcholine antibodies. See Fig. 1 legend for an explanation of the symbols. The sequences of HPC16, HPC35, HPC52 and HPC126 are from ref. 61; M603 is from ref. 64; M167 from ref. 65; T15, W3207, M511, S63, H8 and S107 are from ref. 22; and Y5236 is from J. Hubert (personal communication).

	10 20	HV1 30	Class	T15 Idiotype	Source
HPCM1	D I V M T Q S P T F L A V T A S K K V T I S C	T A S E S L Y S S K H K V H Y	IgM	+	hybridoma
HPCM2	-----------------------	---------------	IgM	+	"
HPCM5*	-----------------------	---------------	IgM	+	"
HPCM6*	-----------------------	---------------	IgM	−	"
HPCM7†	-----------------------	---------------	IgM	−	"
HPC 52	-----------------------	---------------	IgM	+	"
HPCG8*	----------------E------	---------------	IgG3	+	"
HPCG11†	-----------------------	---------------	IgG3	+	"
HPCG12	-----------------------	---------------	IgG3	+	"
HPCG14†	-----------------------	---------------	IgG1	−	"
S63	-----------------------	---------------	IgA	+	myeloma
Y5236	-----------------------	---------------	IgA	+	"
H8	-----------------------	---------------	IgA	+	"
S107	-----------------------	---------------	IgA	+	"
T15	-----------------------	---------------	IgA	+	"
HPCM3	D I V M T Q S P S S L S V S A G E K V T M S C	K S S Q S L L N S G N Q K N Y	IgM	−	hybridoma
HPCM4	-----------------------	---------------	IgM	−	"
HPC126	-----------------------	---------------	IgM	−	"
HPCG15	----S------A-----------	---------RTR---	IgG1	−	"
W3207	-----------------------	-------DG-	IgA	−	myeloma
M603	-----------------R-----	--------------F	IgA	−	"
HPC19	D I V I T Q D E L S N P V T S G E S V S I S C	R S S K S L L Y K D G K T Y L	IgM	−	hybridoma
HPC104	-----------------------	---------------	IgM	−	"
HPC16	--------------V--------	---------------	IgM	−	"
HPCG9†	-----------------------	---------S-----	IgG3	−	"
HPCG10†	-----------------------	----------Q----	IgG3	−	"
HPCG13†	------N----------------	---------------	IgG1	−	"
HPCG17	-----------------------	---------------	IgG1	−	"
M511	----------K------------	---------------	IgA	−	myeloma
M167	-----------------------	---------------	IgA	−	"

from the presence of multiple germ-line D gene segments or from somatic variation. The insertions or deletions of codons associated with either end of the D segment as well as the junctional codon substitutions probably arise from the site-specific recombinational mechanism that joins the V_H and D as well as the D and J_H gene segments (Fig. 4). This diversity arises because the actual DNA splice points may occur at different sites on each of these three gene segments, in each case generating insertions, deletions and/or hybrid codons at the junctional boundaries[6-8]. Figures 3 and 4 suggest that there are probably multiple germ-line D gene segments and that variable boundaries of recombination may occur for any of the V region gene segments—V_H, D or J_H. The blocks of substitutions may arise from multiple germ-line D segments, from the splicing together of two germ-line D gene segments or from a variety of other diversification mechanisms[4,17,36,37].

The number and diversity of germ-line D gene segments are not known, as the D gene segments have only recently been isolated from germ-line DNA (P. Early and L. Hood, personal communication). As a result of D segment variability, the third hypervariable region of the heavy chain is the most diverse portion of the antibody V regions[38]. This can be clearly seen in the V_H regions of immunoglobulins binding phosphorylcholine as well as those binding dextran[3].

The J_H segments are encoded by one germ-line J_H gene segment

The J_H segments of mouse heavy chains are encoded by four different J_H gene segments[8,20]. In the 16 V_H regions from immunoglobulins binding phosphorylcholine that can be compared, all seem to have J_H segments coded by the J_{H1} gene segment (Figs 3, 4). Five J_H segments (HPCM6, HPCM4, HPCG14, M603 and W3207) have N-terminal substitutions or deletions at position 108. As discussed before, this diversity can be explained by variability in the DNA splice point for joining the D and J_H gene segments (Fig. 4). In addition, two J_H segments (W3207 and M167) have additional substitutions when compared with the T15 sequence. At first glance, these data suggest that a particular J_H gene segment may be selected by antibody specificity. However, this selection is not always the rule, for in the 19 V_H regions derived from hybridoma antibodies binding dextran, all four J_H gene segments are expressed[3] (J. Schilling, personal communication).

Antibodies binding phosphorylcholine demonstrate combinatorial association between light and heavy chains

Among the IgM immunoglobulins binding phosphorylcholine, some use the same heavy chain in conjunction with two very different light chains (such as HPCM2 and HPCM3 in Figs 1 and 3). The T15- and M603-like V_L regions associated with these V_H regions differ by more than 50% of their amino acid sequence over their N-terminal 38 residues and belong to two different groups (Fig. 1). Moreover, this combinatorial association of two different light chains with the same heavy chain changes the nature of the antigen-binding site because the affinities of the HPCM2 and HPCM3 antibodies for phosphorylcholine differ by 10-fold (Table 1). It has long been assumed that one heavy chain may pair with many different light chains[39]. Indeed, if 1,000 light and 1,000 heavy chains could be associated combinatorially, 10^6 different antibodies could be produced. Although it has not been formally demonstrated that every light chain can pair with every heavy chain, these data suggest that combinatorial association may occur in some antibodies to increase diversity.

The T15 idiotype is present on immunoglobulins with sequence heterogeneity in V_H and V_L regions

Antisera recognizing the T15 idiotype have been used to identify a restricted group of antibodies that bind to phosphorylcholine. The immunoglobulin class and idiotype classification of antibodies are given in Figs 1–3, and show that antibodies with the T15 idiotype are found among proteins of the IgM, IgG3 and IgA classes. These data also show that antibodies with the T15 idiotype can have different amino acid sequences in their V regions. Four hybridoma antibodies (HPCM1, HPCM2, HPCM5 and HPC52) and four myeloma proteins (S63, Y5236, S107 and T15) exhibit the T15 idiotype and have identical N-terminal V_H and V_L sequences. However, hybridomas HPCG8, HPCG11 and HPCG12 express the T15 idiotype but differ from the proteins described above by one or more residues either in the V_H (HPCG8, HPCG11 and HPCG12) or the V_L (HPCG8) region. These data clearly indicate that antibodies with the same idiotype comprise a

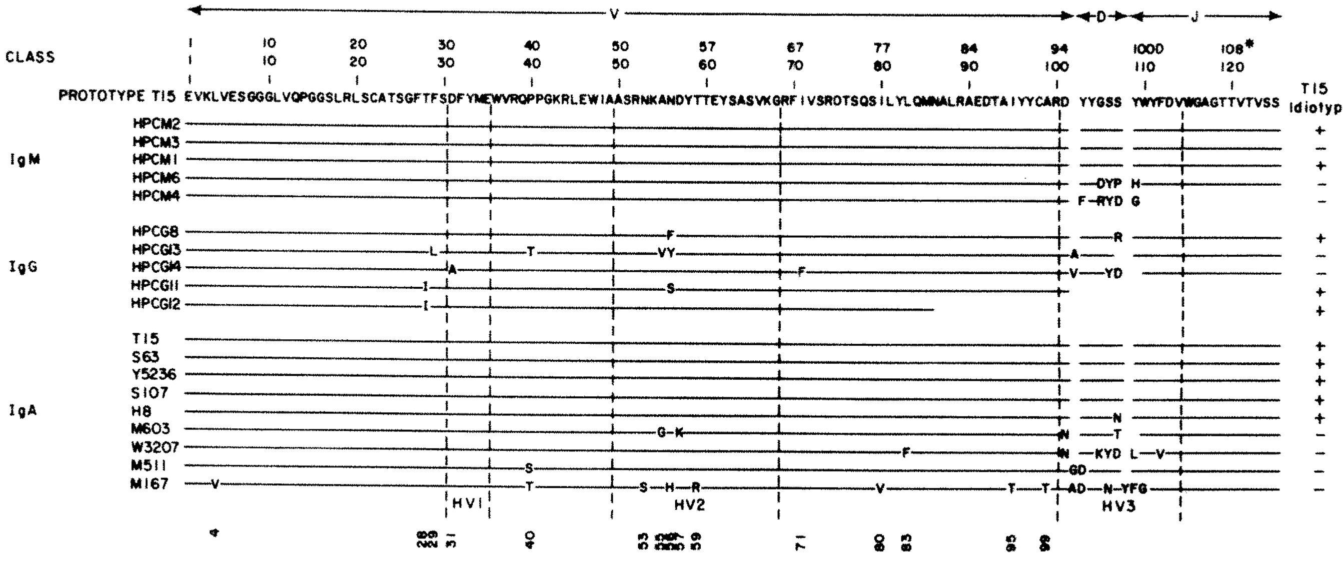

V
D
J
CLASS
PROTOTYPE T15 EVKLVESGGGLVQPGGSLRLSCATSGFTFSDFYMEWVRQPPGKRLEWIAASRNKANDYTTEYSASVKGRFIVSRDTSQSILYLQMNALRAEDTAIYYCARD YYGSS YWYFDVWGAGTTVTVSS
T15 Idiotype
IgM
HPCM2
HPCM3
HPCM1
HPCM6
HPCM4
IgG
HPCG8
HPCG13
HPCG14
HPCG11
HPCG12
IgA
T15
S63
Y5236
S107
H8
M603
W3207
M511
M167
HV1
HV2
HV3

Fig. 3 Protein sequences of the V_H, D and J_H segments of anti-phosphorylcholine antibodies. The strategy for sequence analysis has been presented elsewhere[66]. Hypervariable regions are indicated by vertical dotted lines (HV1, HV2 and HV3)[35] and the V, D and J segment boundaries are shown above the sequences. The numbering system of Kabat *et al.*[35] is shown on the first line(*) and is for M167. The consecutive numbering is shown on the second line and is for the M167 V_H region. All other V_H regions are shorter as indicated by the deletions at the V_H–D and D–J_H boundaries. Positions of amino acid substitutions are labelled at the bottom of the figure except in HV3. The references for data are given in Fig. 2 legend. The Asx residue reported at position 101 in M167 in ref. 35 has been changed to asparagine based on the DNA sequence[8].

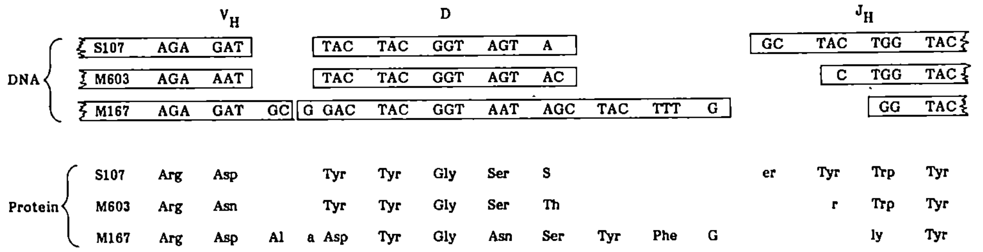

VH
D
JH
DNA
S107 AGA GAT
TAC TAC GGT AGT A
GC TAC TGG TAC
M603 AGA AAT
TAC TAC GGT AGT AC
C TGG TAC
M167 AGA GAT GC
G GAC TAC GGT AAT AGC TAC TTT G
GG TAC
Protein
S107 Arg Asp Tyr Tyr Gly Ser S er Tyr Trp Tyr
M603 Arg Asn Tyr Tyr Gly Ser Th r Trp Tyr
M167 Arg Asp Al a Asp Tyr Gly Asn Ser Tyr Phe G ly Tyr

Fig. 4 Comparison of the DNA and protein sequences of three V_H regions at the V_H-D-J_H boundaries. The three sets of boxes indicate the distinct V_H, D and J_H gene segments. The DNA sequences are from ref. 8.

family of closely related but distinct V regions, and, in this regard, are similar to immunoglobulins which bind nitrophenyl[40], dextran[3] and arsonate[41,42].

Does somatic variation occur in the V_H and V_L segments derived from immunoglobulins binding phosphorylcholine?

The protein sequences presented here demonstrate that the V_H (amino acids 1–101) and V_L (amino acids 1–97) segments are very diverse. The question arises whether this V segment diversity can be generated entirely from multiple germ-line genes, or whether some somatic diversification must occur. We have approached this question by analysing the protein and nucleic acid data for the proteins and genes encoding the phosphorylcholine response.

To determine a minimal estimate of the number of V genes in the T15 V_H group, a cloned V_H probe was used to analyse undifferentiated (embryo) DNA by Southern blot analysis[19]. This V_H probe was derived from a full-length cDNA copy of the heavy chain mRNA from the S107 myeloma tumour[43]. (The S107 and T15 V_H regions have identical protein sequences [Fig. 3].) The S107 V_H probe hybridizes strongly to four restriction fragments. Experiments have demonstrated that the S107 V_H probe cross-reacts strongly with mRNAs from M167 and M603[43]. Because these are among the most different V_H regions from phosphorylcholine-binding immunoglobulins (Fig. 3), the S107 cDNA probe should detect most, if not all, of the V_H genes in the T15 group. If each of the DNA restriction fragments that hybridize to the S107 cDNA probe contains one T15-like V_H gene segment, there are approximately four V_H genes in the T15 V_H group. However, more than four different phosphorylcholine-specific myeloma and hybridoma proteins have already been sequenced. Among the nine myeloma V_H regions that are completely sequenced, four differ in their V segments from the T15 V_H region by one to eight residues (Fig. 3). Among the 20 hybridoma V_H regions that have been examined over their N-terminus, seven differ from the T15 V_H segment and from the myeloma V_H variants (Figs 2, 3). Thus, in the 29 V_H regions derived from immunoglobulins binding phosphorylcholine, 12 different V_H protein segments are present. The four germ-line DNA restriction fragments hybridizing to the S107 V_H probe may each contain more than one V gene segment. However, even if each band contained three V_H gene segments—the number necessary to explain the protein V_H segment data—one would expect, in 29 V_H segments analysed at the protein level, to see several examples of the same germ-line V_H gene being expressed two or three times. Instead, each nonprototype sequence is unique, and with two exceptions, every variant amino acid is different from other amino acids found at that position. Hence, we conclude that the pool of distinct V_H regions is very large and somatic variation must account for at least a portion of the nonprototype V_H sequences reported here. Interestingly, as judged by the data in Figs 2 and 3, variation occurs as frequently outside as within the hypervariable regions (16 of 31 V_H segment substitutions lie outside the hypervariable regions). Therefore, either the framework variants arise from distinct germ-line V_H gene segments or the presumptive mechanism for somatic variation is not confined to the hypervariable regions.

Somatic variation also may occur in anti-phosphorylcholine light chain genes. In the M167 V_L group, there is only one germ-line gene[44] (I. Weissman and U. Storb, personal communication). Six different protein sequences have already been identified (Fig. 1). Thus, somatic diversification also generates V_L variants.

An alternative explanation for this V region diversity is the possibility of genetic polymorphism in V gene segments among BALB/c mice. Although this possibility cannot be ruled out, it seems unlikely. BALB/c mice have been inbred for hundreds of generations, and mice obtained from a single source should be greater than 99% homozygous[45], although admittedly, some polymorphisms do exist in inbred strains[46]. However, the extent of variation seen in the V_H and V_L regions from immunoglobulins binding phosphorylcholine far exceeds any known types of polymorphism including those noted in the globins and haptoglobins. For example, it is striking that four of the V_H and three of the V_L variants analysed in this study are derived from a single mouse (Figs 1, 2). Accordingly, individual mice are capable of generating extensive V segment diversity.

IgG antibodies binding phosphorylcholine exhibit more diversity than their IgM counterparts

The V_H segments from 11 IgM hybridomas do not show any diversity in their N-terminal 36 residues (Fig. 2). In contrast, the corresponding portions of the V_H regions from IgG hybridomas vary in six of the nine examples studied. The same pattern is seen in the corresponding V_L segments (Fig. 1). Of 12 light chains from IgM antibodies, 11 are identical in sequence to the prototype sequence of their respective groups while five of nine V_L regions from IgG antibodies vary. Thus, eight out of nine of the IgG antibodies vary from the V group prototype sequences in either their V_H or V_L regions, whereas only one of the 12 IgM antibodies varies.

One possible explanation for the observed differences in diversification between IgG and IgM proteins is that both do arise from the same population and the differences are the result of sampling errors. We used Fisher's exact method to demonstrate that it is very unlikely that these differences came from sampling error. The probability that the V_L segments of the IgM molecules and the V_L segments of the IgG molecules were selected from the same population is less than 0.05; likewise, the probability that the V_H segments from the IgM and IgG molecules were drawn from the same pool is 0.01; finally, the probability that both the V_L and the V_H segments were drawn from common V_L and V_H populations is less than 0.005. Thus, it is likely that the IgG polypeptides are significantly more diverse than their IgM counterparts. The complete V_H segment sequences strongly support this conclusion (Fig. 3) in that every V_H segment from the IgM group is identical to T15 and every V_H segment from the IgG shows one or more substitutions.

Cellular and recombinant DNA studies have suggested that B cells initially produce IgM and then switch to produce IgG or IgA after antigen stimulation[20,47–49]. Thus, the greater diversity of the IgG hybridomas must arise either from infrequent nonprototype V_H gene segments in IgM-producing B cells, which are selectively expanded and expressed in IgG-producing cells, or from nonprototype V_H gene segments that are generated in

IgG-producing cells. Two important questions arise. First, how and when are the nonprototype sequences generated? Second, what selective pressures could give rise to the differences between IgM and IgG diversity?

The nonprototype sequences could be produced by one of two general mechanisms. (1) The nonprototype V regions may be encoded by germ-line V gene segments. We think this unlikely because there are apparently more protein sequences than germ-line genes in the T15 V_H group and M167 V_L group. (2) The nonprototype V regions may arise by somatic diversification. If so, B cells that express variant antibodies must be amplified by selection after the class switch from IgM to IgG production. Somatic diversification may occur by any one of three distinct mechanisms: ordinary somatic mutation[37], somatic recombination[39], or a special hypermutational mechanism[36,50]. Two points are of special interest with regard to mechanisms of somatic mutation. First, activation of the somatic diversification mechanism may occur in conjunction with the switching of heavy chain classes. If so, the antibodies in secondary immune responses should be far more diverse than their primary immune response counterparts[51]. As the somatic diversification mechanism must generate variants in V_L as well as V_H gene segments, it must be activated in a *trans* fashion on different chromosomes. Second, somatic mutation may occur throughout the lifetime of individual B cells. Accordingly, the oldest B cells (those having undergone most clonal expansion) accumulate the most mutations. To explain our observations by this model, one would suggest that on the average B cells that had undergone a class switch were older than those which had not switched[52].

Two types of selection may operate on infrequent germ-line or somatic variants which could result in the observed IgM–IgG differences. (1) Antigen-driven selection. B-cell clones expressing variant antibodies may be expanded because they have a higher affinity for antigen than B cells expressing prototype antibodies. As IgG-producing cell populations may have undergone more antigen-driven selection than IgM populations, IgG antibodies should show more variability than IgM antibodies. At first sight, the data presented in Table 1 seem to argue against this possibility. The range of affinity constants seen in the IgM and IgG antibodies is not significantly different. However, these data do not exclude the possibility that the affinity constant of antibody for free hapten could be significantly different from the affinity constant of antibody for the hapten–carrier conjugate. Selection for variants in non-binding site regions may occur whenever framework residues modify the positions of binding-site residues. (2) Idiotype selection. Selection of infrequent variants may occur by idiotype-specific regulation. For example, if idiotype-specific suppression controls the level of predominant clones expressing the prototype sequences, B cells with altered sequences may be selectively expanded.

The IgA antibodies also are more diverse than their IgM counterparts ($P<0.05$). However, the IgA V_H regions from immunoglobulins binding phosphorylcholine fall into two categories—those which have diversified (5/9) and those which are identical to the T15 prototype sequence (4/9) (Figs 1, 3). We propose that the V_H and V_L segments of IgG3, IgG1, IgA and probably IgG2a and IgG2b antibodies generally will be significantly more diverse than the V_H and V_L segments of IgM antibodies. Thus, increased diversity seems to be associated with the class switch from IgM to other isotypes. The increased variability in V segments is in part due to the somatic diversification of germ-line V genes. It will be interesting to compare the amino acid sequences of the protein variants with the nucleic acid sequences of the germ-line genes to determine which amino acid residues are encoded in the germ-line DNA and which result from somatic variation.

This work was supported by NIH grants CA 25507 and AI 10781. We thank Nadine Nivera for technical assistance, Fred Karush for help in performing equilibrium dialysis experiments, Julia Hurwitz and Joseph Davie for hybridomas, Doug Fambrough, Henry Huang, Steve Crews and Mitch Kronenberg for helpful discussions, Tim Hunkapiller for computer graphics, and J. Myron Hood for help with the statistical analyses.

Received 18 December 1980; accepted 27 February 1981.

1. Weigert, M. & Riblet, R. *Cold Spring Harb. Symp. quant. Biol.* **41,** 837–846 (1976).
2. Tonegawa, S., Hozumi, N., Mathyssens, G. & Schuller, R. *Cold Spring Harb. Symp. quant. Biol.* **41,** 877–889 (1976).
3. Schilling, J., Clevinger, B., Davie, J. M. & Hood, L. *Nature* **283,** 35–40 (1980).
4. Hood, L. *et al. Cold Spring Harb. Symp. quant. Biol.* (in the press).
5. Weigert, M., Gatmaitan, L., Loh, E., Schilling, J. & Hood, L. *Nature* **276,** 785–790 (1978).
6. Max, E., Seidman, J. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **76,** 3450–3454 (1979).
7. Sakano, H., Hüppi, K., Heinrich, G. & Tonegawa, S. *Nature* **280,** 288–294 (1979).
8. Early, P., Huang, H., Davis, M., Calame, K. & Hood, L. *Cell* **19,** 981–992 (1980).
9. Wu, T. & Kabat, E. *J. exp. Med.* **132,** 211–250 (1970).
10. Capra, J. D. & Kehoe, J. *Proc. natn. Acad. Sci. U.S.A.* **71,** 845–848 (1974).
11. Amzel, L., Poljak, R., Saul, F., Varga, J. & Richard, F. *Proc. natn. Acad. Sci. U.S.A.* **71,** 1427–1430 (1974).
12. Padlan, E., Segal, D., Spande, T., Davies, D., Rudikoff, S. & Potter, M. *Nature new Biol.* **245,** 165–167 (1973).
13. Hood, L., Gray, W., Sanders, B. & Dreyer, W. *Cold Spring Harb. Symp. quant. Biol.* **32,** 133–146 (1967).
14. Milstein, C. *Nature* **216,** 330–332 (1967).
15. Potter, M. *Adv. Immun.* **25,** 141–211 (1977).
16. Bernard, O., Hozumi, N. & Tonegawa, S. *Cell* **15,** 1133–1144 (1978).
17. Seidman, J., Leder, A., Nau, M., Norman, B. & Leder, P. *Science* **202,** 11–17 (1978).
18. Brack, C., Hirama, M., Lenhard-Schuller, R. & Tonegawa, S. *Cell* **15,** 1–14 (1978).
19. Davis, M., Early, P., Calame, K., Livant, D. & Hood, L. *ICN-UCLA Symposium on Eucaryotic Gene Regulation* (eds Axel, R., Maniatis, T. & Fox, C. F.) 393–406 (Academic, New York, 1979).
20. Sakano, H., Maki, R., Kurosawa, Y., Roeder, W. & Tonegawa, S. *Nature* **286,** 676–683 (1980).
21. Köhler, G. & Milstein, C. *Nature* **256,** 495–497 (1975).
22. Hood, L. *et al. Cold Spring Harb. Symp. quant. Biol.* **41,** 817–836 (1976).
23. Kluskens, L., Lee, W. & Köhler, H. *Eur. J. Immun.* **5,** 489–496 (1975).
24. Rudikoff, S. & Claflin, J. L. *J. exp. Med.* **144,** 1294–1304 (1976).
25. Sher, A. & Cohn, M. *Eur. J. Immun.* **2,** 319–326 (1972).
26. Cosenza, H. & Köhler, H. *Science* **176,** 1027–1029 (1972).
27. Sigal, N. H., Gearhart, P. J., Press, J. L. & Klinman, N. R. *Nature* **259,** 51–52 (1976).
28. Lieberman, R., Potter, M., Mushinski, E. B., Humphrey, W. & Rudikoff, S. *J. exp. Med.* **139,** 983–1001 (1974).
29. Gearhart, P. J. & Cebra, J. J. *Nature* **272,** 264–265 (1978).
30. Strayer, D. S., Cosenza, H., Lee, W. M. F., Rowley, D. A. & Köhler, H. *Science* **186,** 640–643 (1974).
31. Bottomly, K. & Mosier, D. E. *J. exp. Med.* **150,** 1399–1409 (1979).
32. Gearhart, P. J. & Cebra, J. J. *J. exp. Med.* **149,** 216–227 (1979).
33. Barstad, P. *et al. Science,* **183,** 962–964 (1974).
34. Claflin, J. L. *Eur. J. Immun.* **6,** 669–674 (1976).
35. Kabat, E. A., Wu, T. T. & Bilofsky, H. *Sequences of Immunoglobulin Chains* (US Department of Health, Education and Welfare, 1979).
36. Baltimore, D. *Nature* **248,** 409–411 (1974).
37. Cohn, M. *et al. ICN-UCLA Symposium on the Immune System: Genes, Receptors, Signals* (eds Sercarz, E., Williamson, A. R. & Fox, C. F.) 89–117 (Academic, New York, 1974).
38. Cebra, J., Koo, P. & Ray, A. *Science* **186,** 263–265 (1974).
39. Gally, J. A. & Edelman, G. A. *Rev. Genet.* **6,** 1–46 (1972).
40. Reth, M., Hämmerling, G. J. & Rajewsky, K. *Eur. J. Immun.* **8,** 393–400 (1978).
41. Marshak-Rothstein, A., Siekevitz, M., Margolies, M. N., Mudgett-Hunter, M. & Gefter, M. L. *Proc. natn. Acad. Sci. U.S.A.* **77,** 1120–1124 (1980).
42. Estess, P., Lamoyi, E., Nisonoff, A. & Capra, J. D. *J. exp. Med.* **151,** 863–875 (1980).
43. Early, P. W., Davis, M. M., Kaback, D. B., Davidson, N. & Hood, L. *Proc. natn. Acad. Sci. U.S.A.* **76,** 857–861 (1979).
44. Joho, R., Weissman, I. L., Early, P., Cole, J. & Hood, L. *Proc. natn. Acad. Sci. U.S.A.* **77,** 1106–1110 (1980).
45. Klein, J. *Biology of the Mouse Histocompatibility-2 Complex* (Springer, New York, 1975).
46. Potter, M. *Cell* **21,** 589–590 (1980).
47. Raff, M. C., Megson, M., Owen, J. J. T. & Cooper, M. D. *Nature* **259,** 224–226 (1976).
48. Davis, M. *et al. Nature* **283,** 733–739 (1980).
49. Kataoka, T., Kawakami, T., Takahashi, N. & Honjo, T. *Proc. natn. Acad. Sci. U.S.A.* **77,** 919–923 (1980).
50. Brenner, S. & Milstein, C. *Nature* **211,** 242–244 (1966).
51. Rodwell, J. D. & Karush, F. *Molec. Immun.* **17,** 1553–1561 (1980).
52. Karush, F. *Comprehensive Immunology*, Vol. 5 (eds Litman, G. & Good, R.) 85–116 (Plenum, New York, 1978).
53. Gearhart, P. J., Sigal, N. H. & Klinman, N. R. *J. exp. Med.* **141,** 56–71 (1975).
54. Shulman, M., Wilde, C. D. & Köhler, G. *Nature* **276,** 269–270 (1978).
55. Kennett, R. H., Denis, K. A., Tung, A. S. & Klinman, N. R. *Curr. Topics Microbiol. Immun.* **81,** 77–91 (1978).
56. Hurwitz, J., Coleclough, C. & Cebra, J. *Cell* **22,** 349–359 (1980).
57. Chesebro, B. & Metzger, H. *Biochemistry* **11,** 766–771 (1972).
58. Johnson, N. D., Hunkapiller, M. W. & Hood, L. E. *Analyt. Biochem.* **100,** 335–338 (1979).
59. Hunkapiller, M. W. & Hood, L. E. *Biochemistry* **17,** 2124–2133 (1978).
60. Dayhoff, M. O. *Atlas of Protein Sequence and Structure* (National Biomedical Research Foundation, Silver Spring, Maryland, 1976).
61. Kocher, H. P., Berek, C., Schreier, M. H., Cosenza, H. & Jaton, J-C. *Eur. J. Immun.* **10,** 264–267 (1980).
62. Barstad, P. *et al. Eur. J. Immun.* **8,** 497–503 (1978).
63. Rudikoff, S. & Potter, M. *Biochemistry* **17,** 2703–2707 (1978).
64. Rudikoff, S. & Potter, M. *Biochemistry* **13,** 4033–4038 (1974).
65. Rudikoff, S. & Potter, M. *Proc. natn. Acad. Sci. U.S.A.* **73,** 2109–2112 (1976).
66. Johnson, N. thesis, California Inst. Technol. (1980).
67. Scatcharo, G. *Ann. N.Y. Acad. Sci.* **51,** 660–672 (1949).
68. Karush, F. *Adv. Immun.* **2,** 1–40 (1962).

Primary structure of a murine transplantation antigen

J. E. Coligan*, T. J. Kindt*, H. Uehara†, J. Martinko† & S. G. Nathenson†

*Laboratory of Immunogenetics, Building 8, Room 100, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20205, USA

† Departments of Microbiology and Immunology and Cell Biology, Albert Einstein College of Medicine, Bronx, New York 10461, USA

A radiochemical technique has been used to determine the primary structure of a murine H–2K^b molecule, a 346-residue glycoprotein subunit attached to a 99-residue β_2-microglobulin subunit. Comparison of this structure with those available for other murine H–2 antigens and human HLA molecules suggests an evolutionary origin of MHC products.

THE antigens encoded within the tightly linked major histocompatibility complex (MHC)[1–4] are a diverse group of membrane-bound molecules intimately involved in various aspects of immune function. This genetic region, which has been characterized in several species, encodes a series of integral cell membrane glycoproteins with a high degree of polymorphism. Although the details of their function remain unknown, it is clear that MHC-encoded molecules are involved in the self–self recognition required for cellular cooperation[1,2] and in the recognition of altered cells by cytolytic T lymphocytes[5,6]. An extensive list of traits, including susceptibility to specific diseases in humans, has also been linked to this complex locus[7,8]. Although the MHC has been studied extensively by immunogenetic methods[3,4], only recently have details of primary structure been elucidated for human and murine MHC gene products[9,10].

Biochemical studies have shown that three major classes of molecules are encoded by the MHC, each having a different role

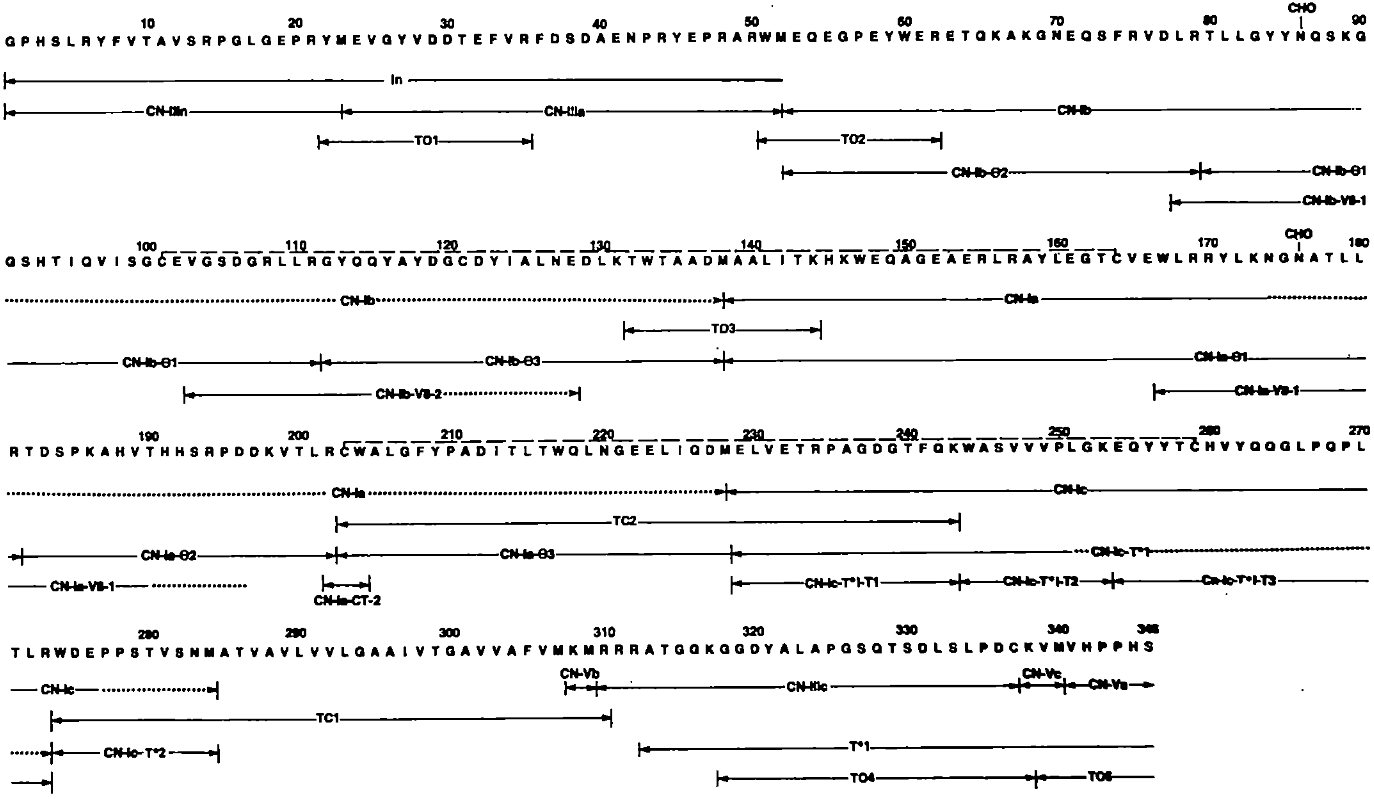

CHO
10
20
30
40
50
60
70
80
90
G P H S L R Y F V T A V S R P G L G E P R Y M E V G Y V D D T E F V R F D S D A E N P R Y E P R A R W M E Q E G P E Y W E R E T Q K A K G N E Q S F R V D L R T L L G Y Y N Q S K G
In
CN-IIIn
CN-IIIa
CN-Ib
TO1
TO2
CN-Ib-Θ2
CN-Ib-Θ1
CN-Ib-V8-1
100
110
120
130
140
150
160
170
CHO
180
Q S H T I Q V I S G C E V G S D G R L L R G Y Q Q Y A Y D G C D Y I A L N E D L K T W T A A D M A A L I T K H K W E Q A G E A E R L R A Y L E G T C V E W L R R Y L K N G N A T L L
CN-Ib
CN-Ia
TO3
CN-Ib-Θ1
CN-Ib-Θ3
CN-Ia-Θ1
CN-Ib-V8-2
CN-Ia-V8-1
190
200
210
220
230
240
250
260
270
R T D S P K A H V T H H S R P E D K V T L R C W A L G F Y P A D I T L T W Q L N G E E L I Q D M E L V E T R P A G D G T F Q K W A S V V V P L G K E Q Y Y T C H V Y H Q G L P E P L
CN-Ia
CN-Ic
TC2
CN-Ia-Θ2
CN-Ia-Θ3
CN-Ic-T*1
CN-Ia-V8-1
CN-Ia-CT-2
CN-Ic-T*I-T1
CN-Ic-T*I-T2
Cn-Ic-T*I-T3
280
290
300
310
320
330
340
346
T L R W D E P P S T V S N M A T V A V L V V L G A A I V T G A V V A F V M K M R R R A T G G K G G D Y A L A P G S Q T S D L S L P D C K V M V H P P H S
CN-Ic
CN-Vb
CN-IIIc
CN-Vc
CN-Va
TC1
CN-Ic-T*2
T*1
TO4
TO5

Fig. 1 Summary of information used to deduce the amino acid sequence of the H–2K^b glycoprotein[12,13,20,21]. The sequence is given in single letter code where A is Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; Y, Tyr. CHO indicates a carbohydrate moiety. Peptides are designated by letters indicating method of chain cleavage (In, intact molecule; CN, cyanogen bromide; TO, tryptic peptide obtained from the intact molecule used to overlap CNBr peptides; T*, trypsin after blockage of Lys with citraconyl groups; Θ, thrombin; V8, staphylococcal V8 protease; CT, chymotrypsin). TC refers to tryptic core, that is, tryptic peptides that were precipitated from the digestion mixture by 5% (v/v) acetic acid, collected by centrifugation, dissolved in 65% acetic acid and isolated by subsequent chromatographic procedures[12]. When multiple designations are used, the order indicates the sequence in which the digestions were performed. For example, TO2 was obtained by tryptic cleavage of the intact chain whereas CN-Ib-V8-2 was produced by V8 treatment of the isolated cyanogen bromide fragment Ib. Fragment CN-Vc results from a reproducible but inappropriate CNBr cleavage. Dotted lines indicate peptide regions for which sequence data were incomplete. The strategy used for determining the complete sequence of the H–2K^b molecule relied on initial isolation of fragments obtained by CNBr cleavage (CN) at Met residues. Subsequent amino acid sequence analyses and use of tryptic overlap (TO) peptides permitted the alignment of these fragments. The only portion of the molecule not recovered as a CNBr fragment was the membrane-associated region (residues 284–307), which was obtained as a tryptic fragment from the intact molecule[56]. The amino acid sequences of the small CNBr fragments IIIn, IIIa, Va, Vb and Vc were determined on the intact fragments. The largest CNBr fragments, Ia and Ib, were reduced to smaller fragments by thrombin cleavage in order to effect their sequence determination. The thrombin fragments were overlapped by isolating fragments obtained by staphylococcal V8 protease digestion and chymotryptic digestion of the appropriate CNBr fragments. Small peptides from CNBr fragment Ic were obtained by trypsinization after blocking of the Lys residues by citraconylation; after removal of the citraconyl groups, the large peptide was redigested with trypsin. Similarly, trypsin digestion after citraconylation was used to obtain fragments for overlapping the COOH-terminal CNBr fragments (IIIc, Vc, Va).

0028-0836/81/190035—05$01.00

in the immune response[7,11]. The class I molecules, which include the classical transplantation antigens, are composed of two noncovalently associated polypeptide chains, a 45,000-molecular weight polymorphic chain encoded by the MHC and a non-MHC-encoded 11,500-molecular weight polypeptide, β_2-microglobulin (β_2M)[9,11]. Studies in our laboratories have concentrated on determining the primary structure of murine class I molecules[12–18], and we report the complete primary structure of one of these, H-2K^b, a molecule which has recently been intensively studied[12,13,18–22]. The membrane-bound heavy chain of this glycoprotein is 346 amino acids long, and has two intra-chain disulphide linkages and two carbohydrate prosthetic groups linked to asparagine residues, whereas the associated β_2M is 99 residues long, contains one intra-chain disulphide bond and is nonglycosylated. All our structural analyses of H-2K^b and several other murine class I antigens have been accomplished by the use of radiochemical methodology[9,23].

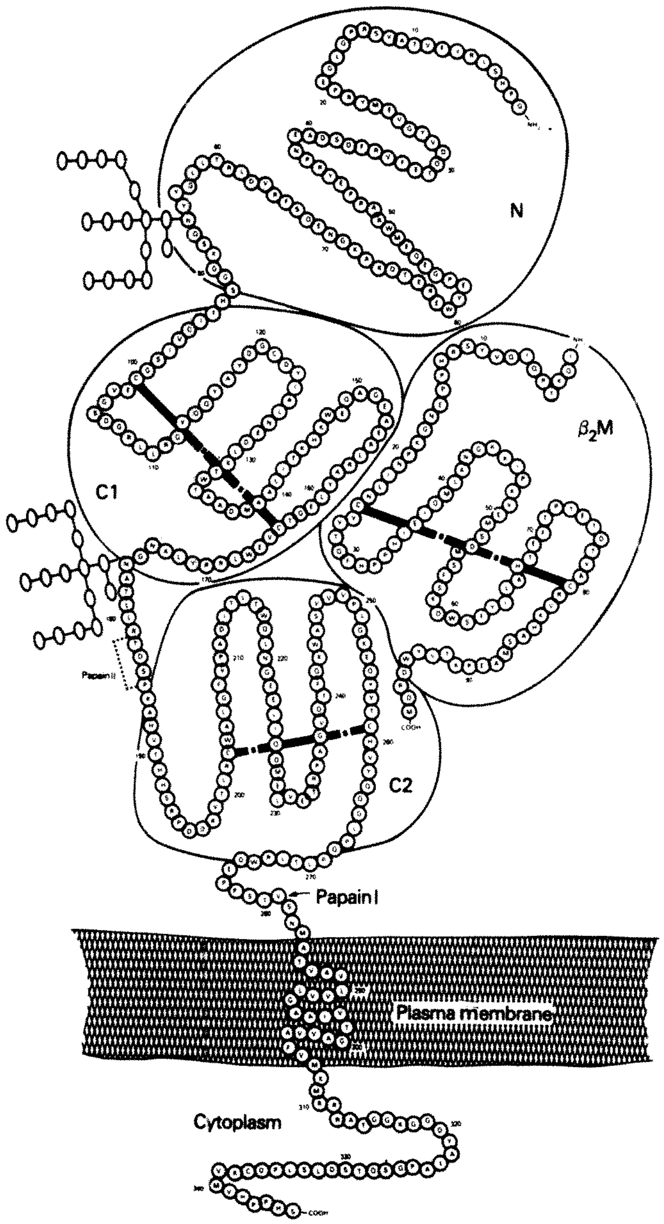


Fig. 2 A schematic representation of the complete amino acid sequence of the murine major transplantation antigen H-2K^b including the associated β_2M and showing putative features of the molecule deduced from primary structure and other properties. The amino acids are identified within the circles in the single letter code detailed in Fig. 1 legend. The Cys–Cys bonds are depicted as black bars. The sequence shown for β_2M is that for the variant found in C57BL/6 mice which has Ala at position 85 (ref. 22). The molecule may be considered to consist of three functional segments: the extracellular segment (~1–281), the membrane-bound segment (~282–308) and the cytoplasmic segment (~309–346). Cleavage by papain at the major site (papain I) yields a soluble antigenically active molecule lacking the hydrophobic and intracellular portions. Division of the extracellular portion of the molecule into three approximately equal subregions is based on homology of the CHO units 90 residues apart (N and Cl) and the presence of two intra-chain disulphide linkages (C1 and C2). The globular nature of the N, C1 and C2 subregions is suggested by limited susceptibility of the molecule to papain digestion; the only fragments yielded are the intact extracellular segments and, in lower yield, a fragment designated papain II that includes the N and C1 subregions[15,57,58]. The dotted box at positions 182–185 shows the approximate location of the papain cleavage yielding papain II. The β_2M is not altered by papain treatment.

Isolation of radiolabelled H-2 alloantigens

There were two major obstacles to the study of the primary structure of H-2 alloantigens. First, because H-2 alloantigens are integral membrane proteins, they possess a hydrophobic region which associates with the membrane lipid bilayer, rendering them insoluble in common aqueous buffers, and thereby preventing the use of conventional biochemical purification techniques. This difficulty has been circumvented by using buffers containing a nonionic detergent (Nonidet P40) which solubilize H-2 antigens in high yield[24]. Second, because only small amounts of H-2 alloantigens can be isolated from tissues[25] or from cultured cells[26,27], our primary structure determinations used H-2 alloantigens biosynthetically labelled with radioactive amino acids[18,28,29].

The procedure for obtaining radiolabelled H-2 antigens has been described previously[14,30]. Briefly, appropriate cells are cultured in the presence of radioactive amino acids and the plasma membranes are then solubilized with Nonidet P40. The detergent-soluble extract is passed over a lentil lectin affinity column to obtain a glycoprotein fraction enriched in H-2 alloantigens. The H-2 molecules are isolated from this fraction by immune precipitation with monospecific alloantisera[17] or monoclonal antibodies[31]. The H-2 heavy chain is dissociated and separated from the light chain (β_2M) by gel filtration in 1 M formic acid.

For radiosequence studies, tumour cell lines are preferred as a source of material because they are metabolically more active than normal cells and thus incorporate greater amounts of the radiolabelled amino acids into cell protein; also tumour cell lines expressing increased amounts of the H-2 molecules under study can usually be obtained[14].

In addition to the increased sensitivity gained in the detection of peptides and phenylthiohydantoin (PTH) amino acids, a major advantage of using radiolabelled material is that the H-2 molecules can then be purified rapidly by immune precipitation[18,28–30] as there is no risk of the radiolabelled H-2 molecules being contaminated with unlabelled proteins (in fact, it is helpful to have additional protein to prevent losses). Also, because the original radiolabelled glycoprotein fraction remains essentially intact, several different molecules can be isolated from the same cell extract by sequential immunoprecipitation steps[31,32].

Sequence determination of radiolabelled proteins

The sequence of radioactive proteins and peptides is determined by stepwise degradation using the automated Beckman sequencer followed by analysis of PTH amino acid derivatives by HPLC, in much the same manner as used for sequencing unlabelled peptides. An obvious difference, however, is that instead of detecting the PTH amino acid derivatives directly on the HPLC, the radioactive PTH derivatives are analysed by co-chromatography with a mixture of unlabelled PTH amino acid derivatives. Peaks are collected and the radioactive amino acid residue is identified by liquid scintillation counting[33].

An important consideration in radiochemical sequence determination concerns the number of amino acids used in a given radiolabelled H-2 preparation. Although this can vary from 1 to 20, difficulties in detecting low specific activity amino acids in the presence of high specific activity amino acids make it practical to use groups of two to six amino acids[14,21]. The radioactive amino acids which become incorporated into cell protein at similar specific activities are grouped together. In determining such groups of amino acids and the amount of each to be incorporated, metabolic interconversions of certain amino acids must be considered[14,33]. For example, radiolabelling cells with glutamine results in radioactive glutamine, glutamic acid and proline in the cell protein in an approximate ratio of 7 : 2 : 1 (ref. 20). This and other interconversions of amino acids have been consistently observed in two murine tumour cell lines[14,33] and a rabbit cell line[34]. Once known, the interconversions present no special difficulty. All amino acids except aspartic acid have been incorporated at sufficient levels for radiosequence analysis. Aspartic acid residues are assigned indirectly after concluding that a given position does not contain one of the other 19 possible amino acids[12,13,21]. (Recent improvements in methods of radiolabelling cells will make possible the direct assignment of Asp residues[35].)

Structural features of H–2K^b

The primary structure of the H–2K^b glycoprotein was determined as outlined in Fig. 1. The primary structure of the β_2M was investigated using radiolabelled protein separated from the H–2K^b, K^d and D^d heavy chains by gel chromatography in 1 M formic acid. Figure 2 depicts schematic representation for the structure of the H–2K^b transplantation antigen suggested from the amino acid sequence and other ancillary data. The 99-residue subunit (β_2M) contains no carbohydrate unit, exhibits limited strain-related polymorphism[22,36–38], is noncovalently associated with the H-2-encoded chain, and shows no indication that it is bound to the cell membrane. The larger of the two chains contains two carbohydrate units and two intra-chain disulphide linkages and exhibits the high degree of polymorphism that characterizes H–2 alloantigens.

A major feature of the molecule is the ease with which the intact extracellular region can be cleaved from the cell surface with the proteolytic enzyme papain[39,40]. Comparison of the structure of the papain-derived molecule (papain I) with that of the intact form has defined the COOH-terminus of the papain-derived molecule as residue 281 (ref. 13). This site coincides with the beginning of a stretch of uncharged and hydrophobic amino acids of approximately 25 residues, thought to be a membrane-binding subregion. (Twenty-five residues in the conformation of an α-helix could span 35 Å, the approximate thickness of the hydrocarbon core of a lipid bilayer[41].) At the carboxy-terminal end of the membrane-binding region there is a cluster of basic amino acids (positions 308 and 310–312) that may serve to anchor the H–2 molecule in the membrane through interaction with the negatively charged phosphate groups of the membrane phospholipids. Glycophorin[42], membrane-bound IgM[43] and HLA[44] have similar clusters of basic residues at the end of the membrane-binding portion of the molecules. That portion of the molecule beginning at the inner face of the membrane and stretching to the C-terminus is apparently located inside the cell[45] and possible functions in the communication of signals to the cell interior following interactions of the exposed extracellular region of the molecule. Whether the phosphoserine residues purported to be present in this portion of H–2[46] and HLA[47] molecules is involved in such processes remains to be demonstrated.

As depicted in the schematic model (Fig. 2), the molecule divides neatly into three functional segments: the extracellular portion which includes the noncovalently bound β_2M and contains the alloantigenic determinants, the membrane segment that penetrates the lipid bilayer and allows the molecule to span the plasma membrane, and the intracellular segment that possibly serves in the communication of signals to the interior of

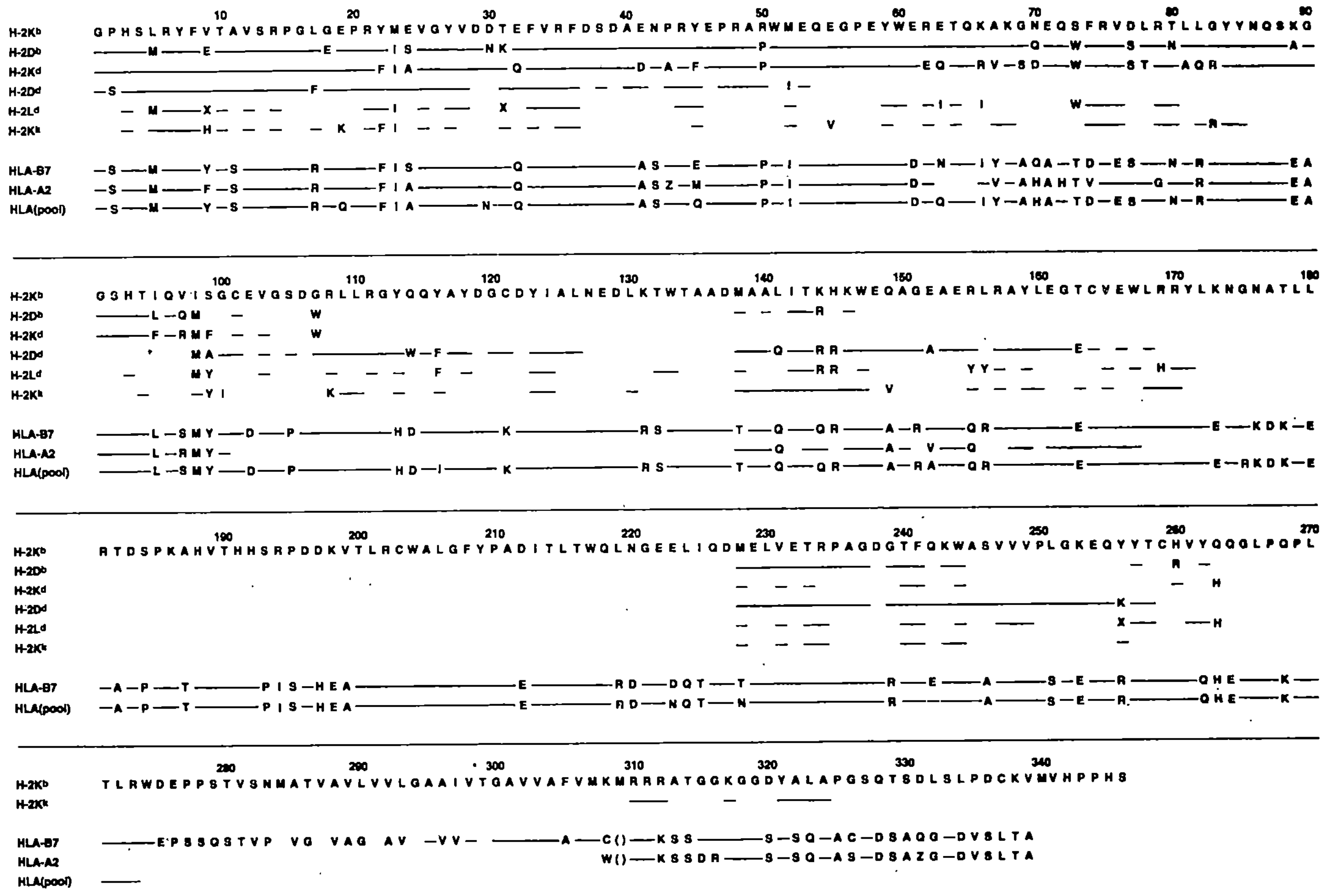

H-2Kb
H-2Db
H-2Kd
H-2Dd
H-2Ld
H-2Kk
HLA-B7
HLA-A2
HLA(pool)
10 20 30 40 50 60 70 80 90
G P H S L R Y F V T A V S R P G L G E P R Y M E V G Y V D D T E F V R F D S D A E N P R Y E P R A R W M E Q E G P E Y W E R E T Q K A K G N E Q S F R V D L R T L L G Y Y N Q S K G
100 110 120 130 140 150 160 170 180
G S H T I Q V I S G C E V G S D G R L L R G Y Q Q Y A Y D G C D Y I A L N E D L K T W T A A D M A A L I T K H K W E Q A G E A E R L R A Y L E G T C V E W L R R Y L K N G N A T L L
190 200 210 220 230 240 250 260 270
R T D S P K A H V T H H S R P D D K V T L R C W A L G F Y P A D I T L T W Q L N G E E L I Q D M E L V E T R P A G D G T F Q K W A S V V V P L G K E Q Y Y T C H V Y Q Q G L P Q P L
280 290 300 310 320 330 340
T L R W D E P P S T V S N M A T V A V L V V L G A A I V T G A V V A F V M K M R R R A T G G K G G D Y A L A P G S Q T S D L S L P D C K V M V H P P H S

Fig. 3 Compilation of amino acid data for histocompatibility antigens for which extensive data are available. The H-2K^b data have been reported in refs 13, 56; H-2D^b, ref. 17; H-2K^d, ref. 16; H-2D^d, ref. 14; H-2L^d, ref. 15; H-2K^k, ref. 46; HLA-B7 and HLA-A2, refs 10, 59, 60; and pooled HLA, ref. 49. Sequences are given in the single letter amino acid code described in Fig. 1 legend. × indicates an unidentified nonidentical residue. Straight lines indicate identity to the K^b sequence. The parentheses denote a position in the HLA molecules where a deletion must be made in order to achieve maximum alignment to H-2K^b.

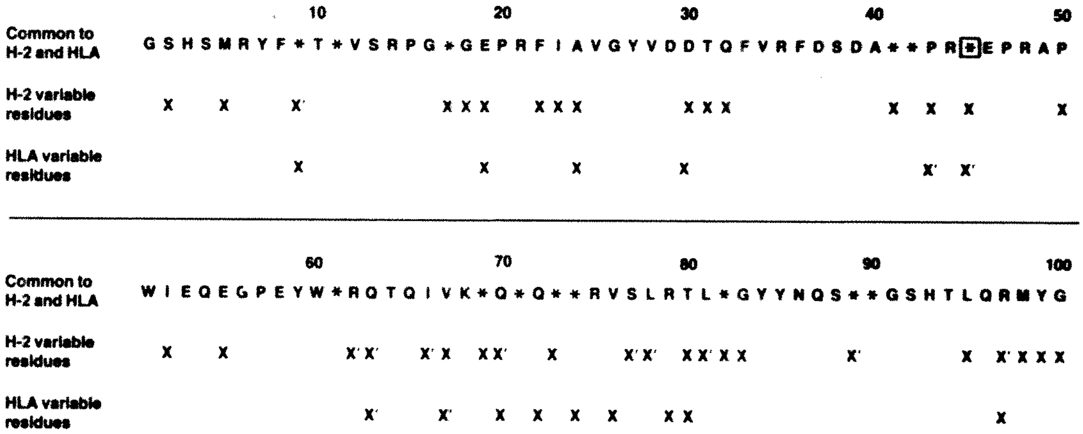

10 20 30 40 50
Common to H-2 and HLA
G S H S M R Y F * T * V S R P G * G E P R F I A V G Y V D D T Q F V R F D S D A * * P R * E P R A P
H-2 variable residues
X X X′ X X X X X X X X X X X X X
HLA variable residues
X X X X X′ X′
60 70 80 90 100
Common to H-2 and HLA
W I E Q E G P E Y W * R Q T Q I V K * Q * Q * * R V S L R T L * G Y Y N Q S * * G S H T L Q R M Y G
H-2 variable residues
X X X′ X′ X′ X X X′ X X′ X′ X X′ X X X′ X X′ X X X
HLA variable residues
X′ X′ X X X X X X X

Fig. 4 Comparison of the H–2 and HLA amino acid sequences for the NH_2 terminal 100 residues. The residues indicated by single letter code (see Fig. 1 legend) are amino acids shared by at least one H–2 and one HLA molecule. At positions where more than one amino acid residue is common to H–2 and HLA, the predominant residue is listed. Asterisks indicate positions where no residues common to H–2 and HLA molecules have been found, but where shared residues could be created by single base changes in the DNA coding sequences. Boxed asterisks are those positions which require a double case change for the generation of shared residues. × indicates positions in H–2 or HLA where multiple substitutions occur but which only require single base changes in the DNA coding sequences for the generation of observed substitutions: ×′ indicates those positions which require multiple base changes for the generation of observed differences.

the cell. Division of the extracellular portion of the molecule into subregions analogous to immunoglobulin-like domains is prompted by primary structural homologies of HLA and β_2M to IgG constant region domains[48,49]. The division (Fig. 2) is supported by the location of the two points where the molecule is susceptible to cleavage by the proteolytic enzyme, papain, which suggest an interruption of a domain-like structure of a globular protein. The subregions (domains) are designated N for the NH_2-terminal portion, C1 and C2 for Cys-containing segments.

Relationship of H-2 and immunoglobulin

The model of the H–2K^b molecule showing domain-like subregions suggests a certain kinship to immunoglobulins. This relationship is supported by weak but significant amino acid sequence homologies between HLA, β_2M and IgG constant regions[48–50], and by the facts that the histocompatibility and immunoglobulin molecules both consist of two chains encoded by unlinked genes and that both the 'heavy' chains are products of complex genetic loci. However, there are some very basic differences. For example, there is no evidence for hypervariable regions or extensive somatic diversification in the H–2K^b molecule as has been found in immunoglobulin variable regions. Whereas pooled immunoglobulin would show heterogeneity in amino acid sequence[51], no evidence of microheterogeneity was found in material isolated from E1–4 tumour cells. Furthermore, when compared by peptide mapping, no differences were observed between H–2K^b isolated from C57BL/6 splenocytes[52] and that isolated from E1–4 tumour cells. In addition, analyses of K^b molecules isolated from splenocytes of mutant mice have revealed sequences identical to those found for the H–2K^b molecules isolated from E1–4[53]. Only discrete differences (one or two amino acids) were found between the H–2K^b molecules isolated from splenocytes from 10 mutant strains of mice and that from E1–4, and they undoubtedly related to the mutations which allowed detection of these molecules in biological assays[53]. Furthermore, examination of the data presented below (Fig. 3) reveals that sequence differences among the H–2 molecules from a given haplotype occur in all parts of the molecule and not just in their NH_2-terminal portions, thus precluding extrapolation of the analogy between H–2 and immunoglobulin molecules to include their having in common structures with clearly delineated variable and constant regions.

Comparison of H–2 and HLA

Although H–2K^b is the only major mouse histocompatibility antigen for which a complete structure is known, extensive areas of several other murine H–2 molecules have been determined and nearly complete primary structures have been reported for the human HLA-B7 and HLA-A2 molecules as well as for a pool of human HLA antigens (Fig. 3). Structural features apparently conserved by all these molecules include the two intra-chain disulphide loops which span positions 101–164 and 203–259 in H–2K^b (ref. 54), and the carbohydrate moiety which is linked to an Asn residue at position 86 (refs 10, 32). Important for the structural studies, which are based on obtaining peptides by CNBr cleavage, is the highly repetitive occurrence of Met residues at identical positions among the different H–2 molecules[32].

Comparison of the amino acid sequence of the NH_2-terminal 100 residues from three H–2, two HLA molecules and the pool of HLA molecules, reveals a high degree of primary structural homology among these molecules. The homology of K^b to D^b, K^b to K^d and K^d to D^b is 85%, 75% and 76%, respectively, and the homology between the HLA molecules is 85%. Comparison between H–2 and HLA molecules yields about 70% homology in all cases. Further examination of the NH_2-terminal sequence data indicates that 63 residues are common to all H–2 molecules (Fig. 3). (These data include the partial sequence data available for K^k, D^d and L^d.) Of these 63 residues, 54 (or 86%) are also common to HLA. Even more remarkably, at 86 of the 100 positions at least one H–2 and one HLA molecule share a common amino acid (Fig. 4). In addition, at only 1 of the 14 positions where common amino acids are not shared would the substitutions observed require a multiple base change in the DNA-coding sequences for common amino acids to occur.

Examination of the NH_2-terminal 100 residues of the H–2 molecules reveals that there are 37 positions where at least two different amino acid residues occur in the sequence; most of these differences are nonrandomly distributed. Small clusters of differences are evident at positions 18–24, 30–32, 41–45 and 95–100, and a large concentration of differences is found between positions 61 and 83, these clusters being interspersed with stretches of conserved amino acid residues (Fig. 4). All these regions represent areas of diversity between the murine and human molecules; however, only the large cluster of differences (61–83) represents an area of diversity when only the human molecules are compared. It is tempting to speculate that the amino acid interchanges observed at positions 61–83 represent an important region of diversity among these alloantigens. In support of this possibility, the homology in this region is strikingly lower than the overall NH_2-terminal homology: only four of the 23 positions (17%) are common to the five molecules for which extensive NH_2-terminal sequence data are available, whereas 50 of the remaining 77 residues (65%) are identical. Likewise, only 10 residues (43%) are shared among the three H–2 molecules in region 61–83 whereas 53 (69%) are shared for the remaining 77 residues. Furthermore, 7 of the 10 amino acid conversions which require two base changes at the nucleic acid level for interconversion occur in this region (Fig. 4).

The extensive sequence data available for H-2K^b and HLA-B7 allow comparisons to be made over the entire length of these molecules. Thus, although the overall homology between these molecules is 71% for positions 1–281 (extracellular portion), the homology for the stretch after position 281 drops to 34%. Comparison of K^b and HLA-A2 in positions available for comparison in this region shows 28% homology. In contrast, the homology between the human molecules (87%)[10] is similar to that observed for the extracellular region of the molecule. In addition to the striking degree of sequence difference between the murine and human molecules, the K^b molecule is 10 residues longer than the human molecules. To achieve maximum alignment of the K^b and B7 molecules, a deletion must be placed at position 307 of the B7 molecule. Also, the second carbohydrate chain present at position 176 in K^b is not detected in HLA-B7.

The finding that there is significantly less homology between K^b and K^d (presumed alleles) than between K^b and D^b is intriguing; it agrees with previous conclusions from peptide mapping data[30] and is difficult to reconcile with the concept that the multiple histocompatibility gene loci arose by gene duplication in a common ancestor. In such a model, divergence of the allelic genes would arise through mutations occurring after speciation. The present data suggest that the gene duplication event occurred subsequent to speciation.

Finally, a major goal of the primary structural studies on H-2 molecules is the localization of sites involved in recognition by alloantibodies and by activated T lymphocytes. It is tempting to speculate that the regions of diversity, especially that found between residues 61 and 83, are involved in such processes. However, although the differences outlined above between the H-2 molecules are quite dramatic, recent data on the structural characterization of H-2 mutants suggest that only small numbers of amino acid differences between antigens are required for immune recognition. Differences between parental and mutant molecules at one or two positions are sufficient to elicit a highly responsive and specific immune recognition, both in alloaggression and in 'associative recognition'[53,55]. Therefore, the multiple differences among the H-2 molecules shown above provide a potentially vast repertoire of recognition determinants for both T-cell allorecognition and serological distinctions.

Parts of this study were supported by NIH grants AI 07289 and AI 10702.

Received 12 January; accepted 18 February 1981.

1. Paul, W. E. & Benacerraf, B. *Science* **195,** 1293–1299 (1977).
2. Klein, J. *Science* **203,** 516–521 (1979).
3. Klein, J. *Biology of the Mouse Histocompatibility-2 Complex* (Springer, New York, 1975).
4. Snell, G. D., Dausset, J. & Nathenson, S. G. *Histocompatibility* (Academic, New York, 1976).
5. Zinkernagel, R. M. & Doherty, P. C. *Adv. Immun.* **27,** 51–177 (1979).
6. Shearer, G. M. & Schmidt-Verhulst, A. M. *Adv. Immun.* **25,** 55–91 (1977).
7. Klein, J. *Adv. Immun.* **26,** 55–146 (1978).
8. Festenstein, H. & Démant, P. *Curr. Topics Immun.* **9** (1979).
9. Nathenson, S. G., Uehara, H., Ewenstein, B., Kindt, T. J. & Coligan, J. E. *A. Rev. Biochem.* (in the press).
10. Strominger, J. L. *et al.* in *The Role of the Major Histocompatibility Complex in Immunobiology* (eds Benacerraf, B. & Cantor, H.) (Garland, New York, 1981).
11. Vitetta, E. S. & Capra, J. D. *Adv. Immun.* **26,** 147–193 (1978).
12. Uehara, H. *et al. Biochemistry* **19,** 6182–6188 (1980).
13. Martinko, J. M. *et al. Biochemistry* **19,** 6188–6193 (1980).
14. Nairn, R., Nathenson, S. G. & Coligan, J. E. *Eur. J. Immun.* **10,** 495–503 (1980).
15. Coligan, J. E. *et al. Proc. natn. Acad. Sci. U.S.A.* **77,** 1134–1138 (1980).
16. Kimball, E. S., Nathenson, S. G. & Coligan, J. E. *Biochemistry* (in the press).
17. Maloy, W. L., Nathenson, S. G. & Coligan, J. E. *J. biol. Chem.* **256,** 2863–2872 (1981).
18. Ewenstein, B. M., Freed, J. H., Mole, L. E. & Nathenson, S. G. *Proc. natn. Acad. Sci. U.S.A.* **73,** 915–918 (1976).
19. Ewenstein, B. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **75,** 2909–2913 (1978).
20. Coligan, J. E. *et al. Molec. Immun.* **16,** 3–8 (1979).
21. Uehara, H. *et al. Biochemistry* **19,** 306–315 (1980).
22. Gates, F. T. III, Coligan, J. E. & Kindt, T. J. *Proc. natn. Acad. Sci. U.S.A.* **78,** 554–558 (1981).
23. Kindt, T. J., Coligan, J. E., Uehara, H. & Nathenson, S. G. *Proc. int. Conf. Chemical Synthesis of Peptides and Proteins* (Elsevier, Amsterdam, in the press).
24. Schwartz, B. D., Kato, K., Cullen, S. E. & Nathenson, S. G. *Biochemistry* **12,** 2157–2164 (1973).
25. Henriksen, O., Robinson, E. A. & Appella, E. *Proc. natn. Acad. Sci. U.S.A.* **75,** 3322–3326 (1978).
26. Rogers, M. J., Robinson, E. A. & Appella, E. *J. biol. Chem.* **254,** 1126–1133 (1979).
27. Freed, J. H., Sears, D. L., Brown, J. L. & Nathenson, S. G. *Molec. Immun.* **16,** 9–21 (1979).
28. Silver, J. & Hood, L. *Nature* **256,** 63–64 (1975).
29. Capra, J. D., Vitetta, E. S., Klapper, D. G., Uhr, J. W. & Klein, J. *Proc. natn. Acad. Sci. U.S.A.* **73,** 3661–3664 (1976).
30. Brown, J. L., Kato, K., Silver, J. & Nathenson, S. G. *Biochemistry* **13,** 3174–3178 (1974).
31. Maloy, W. L., Hämmerling, G., Nathenson, S. G. & Coligan, J. E. *J. immun. Meth.* **37,** 287–299 (1980).
32. Kimball, E. S., Maloy, W. L., Martinko, J. M., Nathenson, S. G. & Coligan, J. E. *Molec. Immun.* **17,** 1283–1291 (1980).
33. Coligan, J. E. *et al. Proc. natn. Acad. Sci. U.S.A.* **75,** 3390–3394 (1978).
34. Kimball, E. S., Coligan, J. E. & Kindt, T. J. *Immunogenetics* **8,** 201–211 (1979).
35. Gates, F. T. III, Kimball, E. S. & Coligan, J. E. (in preparation).
36. Michaelson, J., Rothenberg, E. & Boyse, E. A. *Immunogenetics* **11,** 93–95 (1980).
37. Goding, J. W. & Walker, I. D. *Proc. natn. Acad. Sci. U.S.A.* **77,** 7395–7399 (1980).
38. Robinson, P. J., Graf, L. & Sege, K. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
39. Nathenson, S. G. & Shimada, A. *Transplantation* **6,** 662–663 (1968).
40. Mann, D. L., Rogentine, G. N. Jr, Fahey, J. L. & Nathenson, S. G. *Nature* **217,** 1180–1181 (1968).
41. Dickerson, R. E. & Geis, I. *The Structure and Action of Proteins*, 28 (Harper & Row, London, 1969).
42. Tomita, M., Furthmayr, H. & Marchesi, V. T. *Biochemistry* **17,** 4756–4770 (1978).
43. Rogers, J. *et al. Cell* **20,** 303–312 (1980).
44. Robb, R. J., Terhorst, C. & Strominger, J. L. *J. biol. Chem.* **253,** 5319–5324 (1978).
45. Walsh, F. S. & Crumpton, M. J. *Nature* **269,** 307–311 (1977).
46. Rothbard, J. B., Hopp, T. P., Edelman, G. M. & Cunningham, B. A. *Proc. natn. Acad. Sci. U.S.A.* **77,** 4239–4243 (1980).
47. Pober, J. S., Guild, B. C. & Strominger, J. L. *Proc. natn. Acad. Sci. U.S.A.* **75,** 6002–6004 (1978).
48. Orr, H. T., Lancet, D., Robb, R. J., Lopez de Castro, J. A. & Strominger, J. L. *Nature* **282,** 266–270 (1979).
49. Trägårdh, L., Rask, L., Wiman, K., Fohlman, J. & Peterson, P. A. *Proc. natn. Acad. Sci. U.S.A.* **77,** 1129–1133 (1980).
50. Coligan, J. E. *et al. Transplantn Proc.* (in the press).
51. Kabat, E. A., Wu, T. T. & Bilofsky, H. *Sequences of Immunoglobulin Chains* (NIH Publ. No. 80-2008, 1979).
52. Nisizawa, T., Ewenstein, B. M., Uehara, H., McGovern, D. & Nathenson, S. G. *Immunogenetics* **12,** 33–44 (1981).
53. Nairn, R., Yamaga, K. & Nathenson, S. G. *A. Rev. Genet.* **14,** 241–277 (1980).
54. Martinko, J. M., Halpern, R., Aldersberg, J. R. & Nathenson, S. G. *Molec. Immun.* (in the press).
55. Ewenstein, B. M. *et al. Immunogenetics* **11,** 383–396 (1980).
56. Uehara, H., Coligan, J. E. & Nathenson, S. G. *Biochemistry* (submitted).
57. Nathenson, S. G. *et al.* in *Current Trends in Histocompatibility* (eds Ferrone, S. & Reisfeld, R. A.) (Plenum, New York, in the press).
58. Trägårdh, L., Wiman, K., Rask, L. & Peterson, P. A. *Biochemistry* **18,** 1322–1328 (1979).
59. Orr, H. T., Lopez de Castro, J. A., Parham, P., Ploegh, H. L. & Strominger, J. L. *Proc. natn. Acad. Sci. U.S.A.* **76,** 4395–4399 (1979).
60. Orr, H. T., Lopez de Castro, J. A., Lancet, D. & Strominger, J. L. *Biochemistry* **18,** 5711–5720 (1979).

Major rearrangement in the human β-globin gene cluster

R. W. Jones, J. M. Old, R. J. Trent, J. B. Clegg & D. J. Weatherall

MRC Molecular Haematology Unit, Nuffield Department of Clinical Medicine, John Radcliffe Hospital, Oxford OX3 9DU, UK

A form of $\delta\beta$-thalassaemia results from a major chromosomal rearrangement involving the β-globin gene cluster which includes an inversion of the entire sequence between the ${}^A\gamma$- and δ-globin genes and two deletions.

IN MAN the ε, ${}^G\gamma$-, ${}^A\gamma$-, δ- and β-globin genes that code for the embryonic, fetal and adult β-like globin chains of haemoglobin are closely linked on chromosome 11[1–6]. The hereditary disorders which result in their altered expression include forms of pancellular hereditary persistence of fetal haemoglobin (HPFH) and $\delta\beta$-thalassaemia which are characterized by an absence of δ- and β-chain production and a variable persistence of ${}^G\gamma$-, ${}^A\gamma$-, or ${}^G\gamma$- and ${}^A\gamma$-chain synthesis into adult life.

0028-0836/81/190039—06$01.00

We describe here the analysis of the ε–γ–δ–β gene cluster in DNA of an Indian patient (FB) who has previously been shown by haematological, haemoglobin and family studies to be homozygous for $^{G}\gamma$-$\delta\beta$-thalassaemia[7]. Restriction endonuclease analysis indicates that the disorder results from a previously undescribed chromosomal rearrangement involving two deletions totalling 8.5 kilobases of DNA and an inversion of 15.5 kilobases of DNA in the γ–δ–β gene cluster. Unlike the deletions previously described in $\delta\beta$-thalassaemia[4,8–10] or HPFH[4,11,12], the entire sequence between the $^{A}\gamma$- and δ-globin genes is retained but is inverted. This is the first example of a major inversion as the basis for a human hereditary disease.

Table 1 Size of restriction enzyme fragments

Enzyme	Size in normal DNA (kb)	Size in $^{G}\gamma$-$(\delta\beta)^{\circ}$-thalassaemia DNA (kb)
a Detected with γ-cDNA probe		
*Bam*HI	2.6+5.1+15.5	2.6+5.1
*Bgl*II	13	10.5
*Eco*RI	7.2+1.6+2.7	7.2+1.6+4.2
*Hpa*I	24+5.1+14	24+4.4
*Pst*I	4.0+5.1	4.0+5.1
*Xho*I	5.1	5.1
b Detected with β-cDNA probe		
*Bam*HI	15.5+4.4+1.8+8.1	20
*Bgl*II	7.7+5.0	10.5+6.6
*Eco*RI	2.2+1.7+5.4+3.6	4.2+3.6
*Hpa*I	1.9+1.4+7.0	1.9+17

DNA was isolated from peripheral blood of a patient (FB) homozygous for $^{G}\gamma$-$(\delta\beta)^{\circ}$-thalassaemia and from a normal human spleen by phenol chloroform extraction[29]. DNA samples were digested with restriction endonucleases as indicated and fractionated in a vertical 0.8% agarose gel by electrophoresis for 15 h at 1.5 V cm^{-1}. The DNA fragments were transferred to a nitrocellulose filter by the gel blotting technique of Southern[30] and the nitrocellulose filter hybridized[31] to a cloned γ- or β-cDNA probe labelled with ^{32}P by nick translation[32]. After hybridization at 42 °C for 48 h the filter was washed in 0.1 × SSC at 52 °C as described elsewhere[33] and the bands detected by autoradiography[34]. The γ-cDNA probe was the γ-globin cDNA plasmid JW151 (ref. 35); the β-cDNA probe was β-globin cDNA plasmid JW102 (ref. 35). The 3.6-kilobase *Eco*RI fragment hybridizing with the B-cDNA probe did not have a place in the final structure of FB DNA. This was probably due to hybridization of the β-cDNA probe to the ε gene[1].

An abnormal $^{A}\gamma$ gene, and abnormal δ- or β-globin sequences

Hybridization of several single (Table 1*a*) and double (not shown) restriction enzyme digests of FB DNA to a cloned γ-cDNA probe showed a normal arrangement of restriction sites on the 5′ side of the $^{A}\gamma$ gene up to the *Pst*I site in that gene (A in Fig. 1). Restriction enzymes producing DNA fragments which span this point in normal DNA produced abnormal fragments when used to digest FB DNA, which suggested that DNA on the 3′ side of A (Fig. 1) had been deleted. However, hybridization of a cloned β-cDNA probe to single enzyme digests of FB DNA (Table 1*b*) showed a number of DNA fragments, most of which differed in size from those seen in normal DNA. Among these were a 4.2-kilobase *Eco*RI and a 10.5-kilobase *Bgl*II fragment, both of which were the same size as abnormal fragments found in digests with the same enzymes hybridized to the γ-cDNA probe (Table 1*a*). This suggested that there was one abnormal DNA fragment generated by each enzyme which spanned the remaining 5′ half of the $^{A}\gamma$ gene and also included some part of either the δ or β genes.

To test this hypothesis, the 4.2-kilobase *Eco*RI and 10.5-kilobase *Bgl*II fragments were digested with *Xho*I. The latter cuts once in the $^{G}\gamma$ gene and once in the $^{A}\gamma$ gene very close to A (Fig. 1) but nowhere else in the whole ε–γ–δ–β region. Thus the

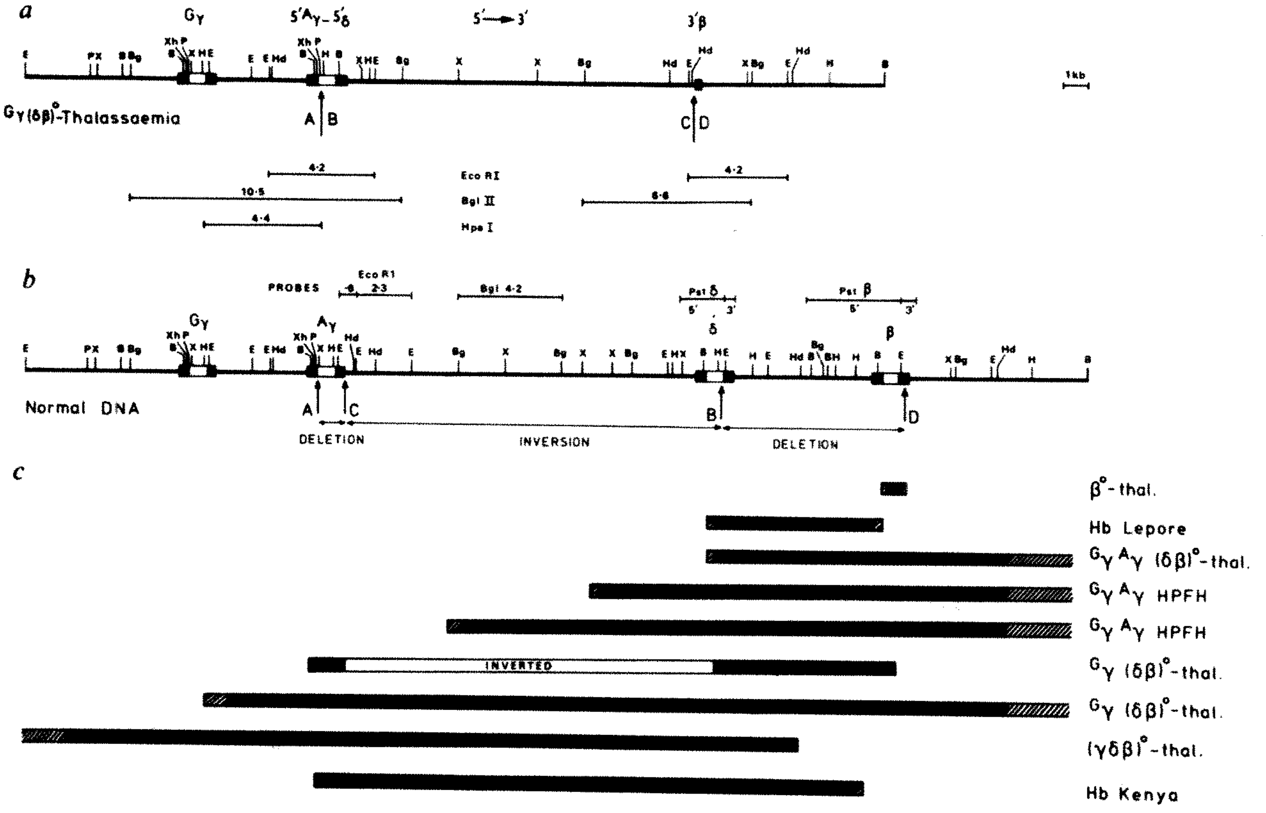


Fig. 1 A map of restriction endonuclease sites in and around the β-like globin gene region of $^{G}\gamma$-$(\delta\beta)^{\circ}$-thalassaemia DNA compared with normal DNA. *a*, Restriction map of DNA from a patient with $^{G}\gamma$-$(\delta\beta)^{\circ}$-thalassaemia. The brackets below the map indicate the locations of some of the abnormal fragments listed in Table 1. *b*, Restriction map of normal DNA derived from refs 4, 24 and 25, except for the *Hin*dIII site on the 5′ side of the δ-globin gene, which was located by hybridization of the 2.3-kilobase *Eco*RI fragment to FB DNA (Fig. 5) and to normal DNA (data not shown). A, B, C and D mark the boundaries of the normal sequences which have been deleted and inverted in the $^{G}\gamma$-$(\delta\beta)^{\circ}$-thalassaemia DNA. The brackets above the normal DNA map indicate the positions of the *Eco*RI 0.6, *Eco*RI 2.3, *Bgl*II 4.2, *Pst* δ and *Pst* β fragments. *c*, Schematic map of the deletions in the β-like globin region in various diseases. The deletions are shown by solid boxes and the end points are located within the hatched boxes where known. The maps are from the following refs: β°-thalassaemia, ref. 26; Hb Lepore, 2,27; $^{G}\gamma^{A}\gamma$-HPFH, 4,11,12; $^{G}\gamma^{A}\gamma$-$(\delta\beta)^{\circ}$ thalassaemia, 4,8,9; $^{G}\gamma$-$(\delta\beta)^{\circ}$ thalassaemia, 4,10; $(\gamma\delta\beta)^{\circ}$-thalassaemia, 21. Different locations of the 3′ end of the $^{G}\gamma^{A}\gamma$-$(\delta\beta)^{\circ}$-thalassaemia deletion have been proposed[4,8,9] and its precise location is unknown. Hb Kenya has not been proved to be a deletion by restriction enzyme mapping. Restriction enzyme sites are: B, *Bam*HI; Bg, *Bgl*II; E, *Eco*RI; Hd, *Hin*dIII; H, *Hpa*I; P, *Pst*I; X, *Xba*I; and Xh, *Xho*I.

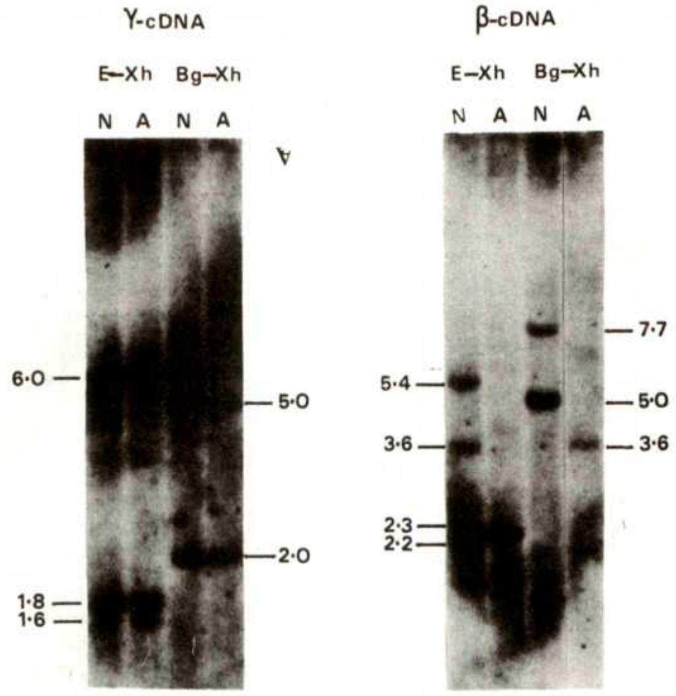


Fig. 2 Hybridization of double digests of normal and FB DNA to cloned γ- and β-cDNA probes. Normal (N) and FB DNAs (A) were digested with either *Xho*I and *Eco*RI (E-Xh) or with *Bgl*II and *Xho*I (Bg-Xh), and duplicate filters were hybridized to cloned γ- and β-cDNAs as described in Table 1 legend. Although a proportion of *Xho*I sites in DNA from some tissues are modified by methylation and therefore are resistant to digestion, there was no evidence of this in the DNA digested here. The size in kilobases of each fragment is shown.

4.2-kilobase *Eco*RI fragment should be cleaved by *Xho*I to give one γ gene fragment and one β-like gene fragment. Double digests of FB DNA with *Eco*RI and *Xho*I (Fig. 2) generated a 1.8-kilobase fragment which hybridized only to the γ-cDNA probe and a 2.3-kilobase fragment which hybridized only to the β-cDNA probe. The sum of these two fragments is 4.1 kilobases, which is close to the expected sum of 4.2. Similarly, *Xho*I cut the *Bgl*II 10.5-kilobase fragment twice to give fragments of 5.0 and 2.0 kilobases which hybridized to γ-cDNA, but only one fragment of 3.6 kilobases on hybridization with β-cDNA (Fig. 2). The sum of these three fragments is 10.9 kilobases, again in close agreement to the predicted value of 10.5.

Both these double digests can therefore be interpreted as showing that in FB DNA, the normal $^{A}\gamma$-gene sequence changes to one containing either all or part of either the β or δ gene, or parts of both. Some β-like sequence must lie within 2.3 kilobases of the *Xho*I site in the $^{A}\gamma$ gene to account for the presence of a DNA fragment of that size in *Eco*RI/*Xho*I digests hybridized to cloned β-cDNA probe.

Inversion of DNA sequence

The relative positions of several new restriction sites, including those of *Bgl*II and *Eco*RI, appeared to be the reverse of those on the 5′ side of the δ-globin gene in normal DNA. An inversion of the normal DNA sequence between A and B (Fig. 1*b*) would place the 5′ half of the δ gene next to the 5′ half of the $^{A}\gamma$ gene, and generate 3.9-kilobase *Eco*RI and 10.7-kilobase *Bgl*II fragments which would hybridize to both γ- and β-cDNA probes (Fig. 1*a*). To determine which β-like sequence lay close to A, 5′- and 3′-δ and β probes were hybridized to *Eco*RI and *Bgl*II digests (Fig. 3). Neither the 3′-δ nor 5′-β probes hybridized to FB DNA, suggesting that these DNA sequences had been lost in the rearrangement. The 5′-δ probe hybridized to 4.2-kilobase *Eco*RI and 10.5-kilobase *Bgl*II fragments supporting the presence of the 5′-δ sequence close to the 5′ half of the $^{A}\gamma$ gene. However, the 3′-β probe hybridized to a 4.2-kilobase *Eco*RI fragment as well. As it also hybridized to a 6.6-kilobase *Bgl*II fragment, rather than one of 10.5 kilobases, it seemed likely that the 3′-β sequence was involved in some further rearrangement of the normal DNA sequence such as the point of reattachment of an inverted sequence to the chromosome (junction C/D, Fig. 1*a*).

The *Eco*RI 2.3- and 0.6-kilobase probes are generated from the normal DNA sequence close to the 3′ end of the $^{A}\gamma$ gene (Fig. 1*b*). Both these probes, as well as the 3′-β probe, hybridized to the 6.6-kilobase *Bgl*II fragment from FB DNA (Fig. 4), showing that this fragment contains sequences complementary to those at both C and D in Fig. 1*b*. The 4.2-kilobase *Eco*RI fragment that hybridized with the 5′-δ probe contains no *Bgl*II site, distinguishing it from the *Eco*RI fragment of the same size to which both the *Eco*RI 0.6-kilobase and 3′-*Pst* β probes hybridized (Fig. 4).

This confirms the separate identities of the two abnormal 4.2-kilobase *Eco*RI fragments, and demonstrates that the relative positions of *Eco*RI and *Bgl*II restriction sites around the junction C/D (Fig. 1*a*) are consistent with the proposed DNA rearrangement.

Linkage of γ, δ and β sequences

To establish the overall arrangement of the sequences in FB DNA, single and double digests with *Hin*dIII and *Xba*I were hybridized to a series of probes complementary to DNA sequences spanning the $^{A}\gamma$ and δ genes (Fig. 5). Hybridization of these digests to the γ–cDNA probe supported the normal location of *Hin*dIII and *Xba*I restriction sites on the 5′ side of the $^{A}\gamma$ gene. The same digests also demonstrated the presence of abnormal 7.2-kilobase *Xba*I and 18.5-kilobase *Hin*dIII fragments which must include the 5′ half of the $^{A}\gamma$ gene, and extend

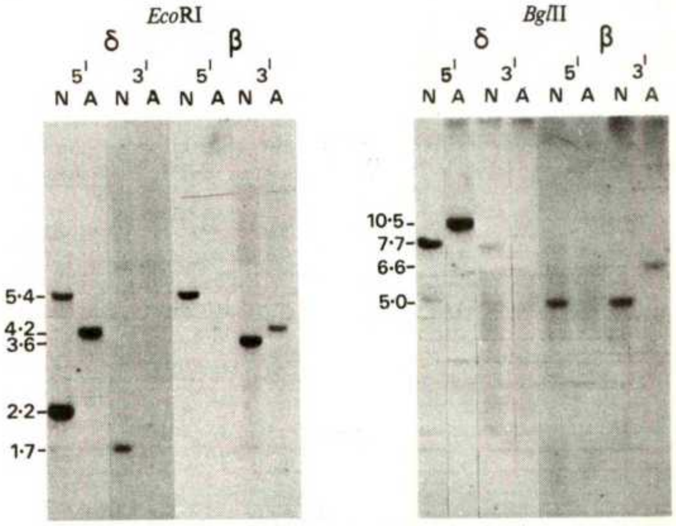


Fig. 3 Hybridization of *Eco*RI and *Bgl*II digests of normal and $^{G}\gamma$-($\delta\beta$)-thalassaemia DNA to 5′- and 3′-*Pst* δ and 5′- and 3′-*Pst* β probes. Normal and FB DNAs were digested with either *Eco*RI or *Bgl*II and duplicate filters hybridized with each of these probes as described in Table 1 except that the hybridization buffer contained 10% dextran sulphate; hybridization was for 12 h, and the filters were finally washed in 0.1 × SSC at 65 °C (ref. 28). The probes were prepared by agarose gel electrophoresis of *Eco*RI digests of the cloned *Pst* β or *Pst* δ fragments. The DNA fragments containing the sequence on either the 5′ or 3′ side of the *Eco*RI site within the δ-globin gene and the β-globin gene were identified after staining with ethidium bromide and eluted electrophoretically from the agarose gel. The size in kilobases of each fragment is shown. *Eco*RI cuts inside the normal δ-globin gene to give a 2.2-kilobase fragment containing 5′ sequences and a 1.7-kilobase fragment containing 3′ sequences; inside the normal β-globin gene it gives a 5.4-kilobase fragment containing 5′ sequences and a 3.6-kilobase fragment containing 3′ sequences. *Bgl*II cuts outside the normal δ- and β-globin genes to give a 7.7-kilobase fragment containing both 5′ and 3′ δ-globin sequences and a 5.0-kilobase fragment containing both 5′ and 3′ β-globin sequences. In *Eco*RI digests of normal DNA there is no significant cross-hybridization of the 5′ probes to 3′ sequences, or vice versa. There is limited cross-hybridization between two 5′ probes, and between the two 3′ probes, but the absence of any significant hybridization of the 5′-*Pst* β or 3′-*Pst* δ probes to FB DNA eases any problems of interpretation. N, normal DNA; A, FB DNA; δ, *Pst* δ probe; β, *Pst* β probe.

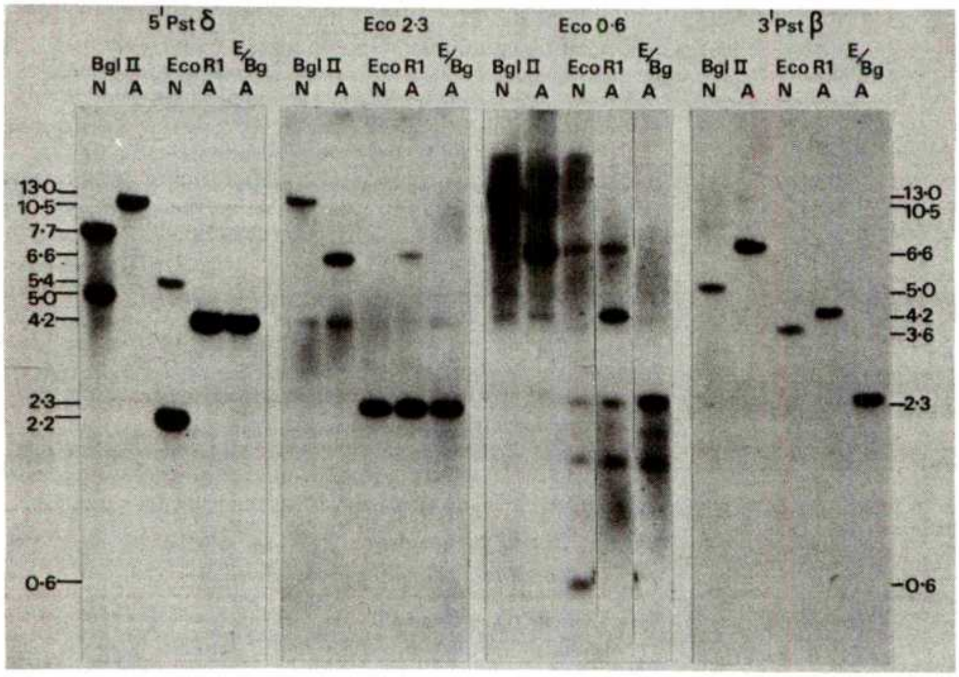


Fig. 4 Hybridization of FB DNA with 5'-*Pst* δ, *Eco*RI 2.3, *Eco*RI 0.6 and 3'-*Pst* β probes. The 3'- and 5'-*Pst* probes were prepared as described in Fig. 3, and the *Eco*RI 0.6 and 2.3 probes in the same way after digestion of λ HγG4[5] with *Eco*RI. The normal (N) and FB (A) DNAs were digested with *Eco*RI, *Bgl*II, or with both enzymes (E/Bg), and duplicate filters were hybridized to each of these probes as described in Fig. 3 legend. The sizes of DNA fragments in kilobases are indicated. The *Eco*RI 0.6-kilobase probe hybridizes to part of the 3' coding sequence of both γ genes. Thus in *Eco*RI digests of normal and FB DNA, it hybridizes to a 1.6-kilobase fragment containing the 3' end of the $^{G}\gamma$ gene as well as to a 0.6-kilobase fragment containing the 3' end of the $^{A}\gamma$ gene (normal DNA) and a 4.2-kilobase fragment (FB DNA). In *Bgl*II digests, as well as hybridizing to a 6.6-kilobase fragment in FB DNA, this probe hybridized to the abnormal 0.5-kilobase fragment containing the intact $^{G}\gamma$ gene. In addition, cross-hybridization occured with a number of unidentified fragments present in both normal and FB DNA.

downstream from it. Both of these fragments were seen when these digests were hybridized with the 5'-*Pst* δ probe, but only the 18.5-kilobase *Hin*dIII fragment also hybridized with the 4.2-kilobase *Bgl*II and 2.3-kilobase *Eco*RI probes (Fig. 5). The sequences hybridizing to these two probes must therefore lie downstream from the $^{A}\gamma$ and δ sequences. Because the 2.3 *Eco*RI probe hybridized with a 0.9-kilobase *Hin*dIII fragment, as well as with the 18.5-kilobase *Hin*dIII fragment, the sequence hybridizing with this probe must straddle the *Hin*dIII site at the 3' end of the large *Hin*dIII fragment. As we already know the relative orientation of the 5'-$^{A}\gamma$ and 5'-δ sequences, we can order the sequences hybridizing to the 18.5-kilobase *Hin*dIII fragment as: 5'-$^{A}\gamma$; 5'-*Pst* δ; 4.2-kilobase *Bgl*II; 2.3-kilobase *Eco*RI.

In digests with *Xba*I, the 4.2-kilobase *Bgl*II, 2.3-kilobase *Eco*RI and 3'-*Pst* β probes all hybridized to an abnormal 9.0-kilobase *Xba*I fragment. Double digests with *Xba*I and *Hin*dIII produced a 5.6-kilobase fragment which hybridized with the first two probes, but not with the 3'-*Pst* β probe. Therefore the 9-kilobase *Xba*I fragment overlaps the 3' end of the 18.5-kilobase *Hin*dIII fragment by 5.6 kilobases, and the 3'-*Pst* β probe must hybridize to a sequence downstream from this overlap, on the 3' side of the sequence hybridizing with the 2.3-kilobase *Eco*RI probe.

More precise information about the relative positions of the sequences hybridizing to the 4.2-kilobase *Bgl*II, 2.3-kilobase *Eco*RI and 3'-*Pst* β probes was provided by these single and double digests. The *Bgl*II probe hybridized to a 2.0–kilobase *Xba*I fragment as well as to the 9.0–kilobase *Xba*I fragment, showing that this sequence straddled the 5' end of the latter fragment. The sequence hybridizing to the 2.3–kilobase *Eco*RI probe must lie near the middle of the 9.0-kilobase *Xba*I fragment straddling one of the two *Hin*dIII sites. The 3'-*Pst* β probe hybridized with a 4.7–kilobase *Hin*dIII fragment which, as shown by a *Hin*dIII–*Xba*I double digest, overlaps the 3' end of the 9.0-kilobase *Xba*I fragment by 2.4 kilobases, with all the sequence hybridizing with this probe contained in this last fragment.

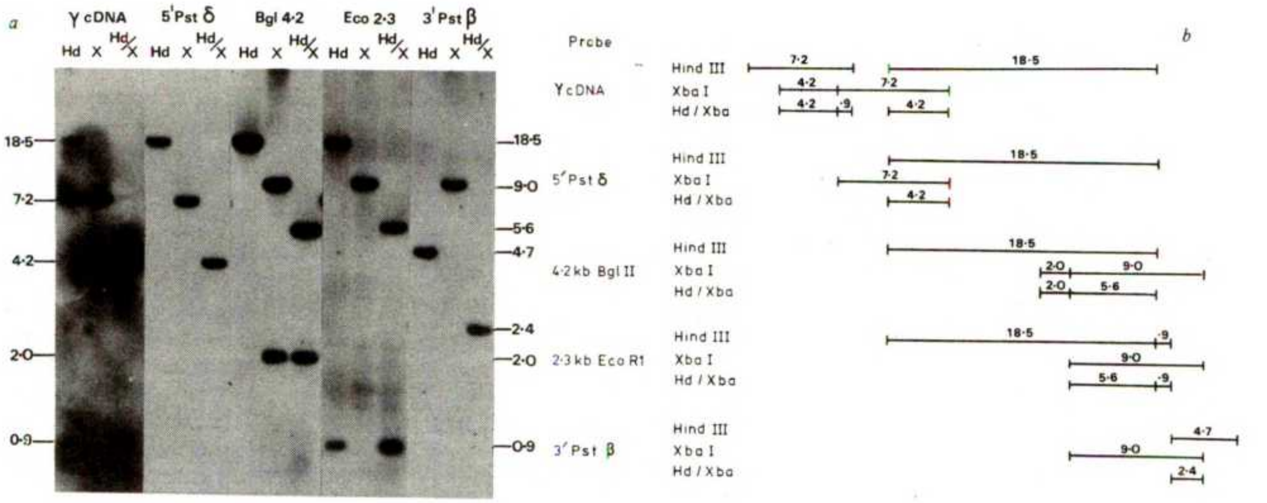


Fig. 5 *a*, Hybridization of FB DNA with the cloned γ-cDNA, 5'-*Pst* δ, 4.2 *Bgl*II, 2.3 *Eco*RI and 3'-*Pst* β probes. The 5'-*Pst* δ, 3'-*Pst* β and 2.3 *Eco*RI probes were prepared as described in the legends to Figs 3 and 4, and the 4.2-kilobase *Bgl*II fragment was prepared in the same way after digestion of HγG4 (ref. 5) with *Bgl*II. FB DNA was digested with *Hin*dIII (Hd), *Xba*I (X), or with both enzymes (Hd-X) and duplicate filters hybridized to each of these probes as described in Fig. 3 legend. *b*, The interpretation of the hybridization data is shown as a series of overlapping fragments spanning the region between the 5' end of the $^{A}\gamma$ gene and the 3' β sequence (see Fig. 1). The size of each fragment is shown in kilobases (kb).

The order of these sequences in FB DNA is therefore 5'-$^{A}\gamma$; 5'-*Pst* δ; 4.2-kilobase *Bgl*II; 2.3-kilobase *Eco*RI; 3'-*Pst* β. In this context it is unlikely that the 0.6-kilobase *Eco*RI probe (Fig. 4) which is contiguous with the 2.3-kilobase *Eco*RI probe in normal DNA, lies anywhere else but between the 2.3-kilobase *Eco*RI and 3'-β sequences. The relative position of these sequences, and of the *Xba*I and *Hin*dIII sites, in this region of FB DNA is entirely consistent with the proposed structure (Fig. 1*a*).

The completed map allows the points A, B, C and D (Fig. 1) to be localized quite precisely; A lies between *Pst*I and *Xba*I sites near the 5' end of the major intervening sequence in the $^{A}\gamma$ gene, and B is between *Hpa*I and *Eco*RI sites at the 3' end of the δ major intervening sequence. C lies in a 0.6-kilobase *Eco*RI fragment at the 3' end of the $^{A}\gamma$ gene, and, as a cloned β-cDNA probe detected abnormal fragments containing the 3' end of the β gene, D must lie between an *Eco*RI site in the 3' coding sequence of the β gene and the end of the gene.

Mechanism of rearrangement of the non-α gene cluster

The form of δβ-thalassaemia described here seems to result from a major rearrangement of the non-α globin gene cluster which has involved the inversion of 15.5 kilobases of DNA between the $^{A}\gamma$ and δ genes and deletions of approximately 1 kilobase from the $^{A}\gamma$ gene and 7.5 kilobases from between the δ and β genes. This is quite unlike previous findings for δβ-thalassaemia[4,8-10] and HPFH[4,8,11,12] which have been shown to result from simple deletions of varying sizes. The rearrangement found in the present case must have resulted from a more complicated mechanism.

It has been suggested that inversions are created between inverted repeat sequences[13-15]. The present arrangement must have involved at least one inversion, but computer analysis of the known sequences close to points A, B, C and D[16-18] (Fig. 1) does not reveal any extensive inverted repeat sequences. Nevertheless, the most economical models for the generation of this rearrangement exclusively involve recombination events of the same kind as would occur between one pair of inverted repeats at A and B, and between a second pair at C and D.

In the case of an intrachromosomal rearrangement, a cross-over between A and B (Fig. 1*b*) in the $^{A}\gamma$ and δ genes of a single chromosome leads to the inversion of the sequence between these points. This inversion brings the sequence between A and C into continuity with that between B and D, and in the process inverts the repeat sequence at C. C and D are now direct repeats and a second cross-over between C and D deletes the sequences between A and C and B and D, producing the abnormal sequence found here. Alternatively, an inversion between C and D is followed by a deletion between A and B. In this model, both cross-overs are of the same kind and each could have occurred independently giving rise to an inversion within a monocentric chromosome.

An alternative explanation is that the cross-overs were between two chromosomes. The repeat sequences at A and D in the $^{A}\gamma$ and β genes of one chromosome are aligned with those at B and C, respectively, in the δ and $^{A}\gamma$ genes of a second. A cross-over between either A and B or D and C leads to an inversion with the products consisting of one bicentric and one acentric chromosome, preventing the formation of viable daughter cells. A second cross-over between the second pair of inverted repeat sequences is therefore obligatory, generating two monocentric chromosomes, one with the arrangement found here.

Despite their economy, these intra- and inter-chromosomal mechanisms still pose problems. Whatever the probability of a single cross-over, the likelihood of two occurring within 20 kilobases of DNA must be smaller. In the case of two cross-overs between chromosomes the probability is even less because they must both be of the same kind and occur at the same time. One solution is that both cross-overs were closely linked as coincident events in a single chromosome (Fig. 6). In this model both cross-overs involve a single $^{A}\gamma$ gene, and are therefore separated by no more than 1 kilobase of DNA. This degree of proximity might allow either cross-over to promote the second, or conceivably both may have been part of a single process—this would account for the fact that they are of the same kind.

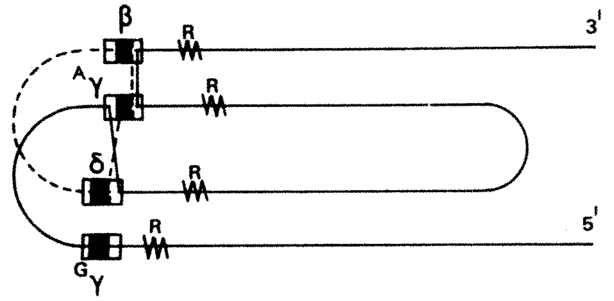


Fig. 6 An intrachromosomal mechanism for the generation of an inversion and deletion. The γ-, δ- and β-globin genes on a single chromosome are aligned so that the cross-overs necessary to generate the inversion and deletion are facilitated. The alignment is brought about by arranging the DNA in a series of loops. The repeat sequences[5] (R) which lie on either side of the pair of γ genes, and at similar distances on either side of the δ and β genes, might interact and serve to stabilize this looped structure. The distances between the repeats and the nearest gene on the left are to scale. If either cross-over occurs alone, an inversion results. When both occur in conjunction, an inversion and deletion as described in this case of $^{G}\gamma$-$(\delta\beta)^{\circ}$-thalassaemia (continuous line) and an acentric fragment (broken line) are formed.

Regulation of γ- and β-chain synthesis

δβ-Thalassaemia and HPFH constitute a spectrum of conditions in which there is significant γ-chain synthesis in adult life[19]; in δβ-thalassaemia this is insufficient to compensate for the absence of δ- and β-chain production, whereas in HPFH γ-chain synthesis is more efficient so that there is less imbalance of globin-chain production.

The deletion which gives rise to $^{G}\gamma^{A}\gamma$-δβ-thalassaemia leaves the 5' end of the δ-globin gene intact whereas the differently sized deletions that produce $^{G}\gamma^{A}\gamma$-HPFH all include the 5' end of the δ-globin gene and variable regions upstream from it (Fig. 1*c*). It has been suggested, therefore, that there may be a region of DNA upstream from the δ-globin gene which is involved in the regulation of the switch from γ- to β-chain synthesis[4,8-12,20]. Loss of this region in the various deletions which produce HPFH might result in failure of the neonatal switch from γ- to β-chain production and hence persistent activity of the $^{G}\gamma$ and $^{A}\gamma$ loci at a level approaching that which occurs in fetal life. In the case of $^{G}\gamma$-$(\delta\beta)^{\circ}$-thalassaemia, a single γ-globin gene does not compensate completely for absent δ- and β-chain production, but there is a high level of $^{G}\gamma$-chain synthesis which must be related to the underlying molecular lesion.

In the one previously reported case of $^{G}\gamma$-$(\delta\beta)^{\circ}$-thalassaemia that has been characterized by restriction enzyme analysis there is a deletion involving all the sequences downstream from a point between the $^{G}\gamma$ and $^{A}\gamma$ genes[4,10] (Fig. 1*c*). The phenotypes of this and the present case of $^{G}\gamma$-$(\delta\beta)^{\circ}$-thalassaemia are very similar but of the sequences missing from the first, those between the $^{A}\gamma$ and δ genes, and downstream from the β gene, are intact in our case. If the loss of putative regulatory sequences in either of these regions is responsible for persistent $^{G}\gamma$-globin synthesis in the first case of $^{G}\gamma$-$(\delta\beta)^{\circ}$-thalassaemia, then the inversion of the sequence between the $^{A}\gamma$ and δ genes, or the interposition of this rearranged sequence between the $^{G}\gamma$ gene and the region downstream from the β gene, must have the same effect in ours. Any regulatory sequences in these regions must therefore be inactivated by the rearrangement of DNA around them, which implies that they are acted on, or act, in *cis*. This comparison does not rule out the existence of regulatory sequences in these regions, but it does place constraints on their nature.

The observation that an intact β-globin gene is inactive in a case of $(\gamma\delta\beta)^{\circ}$-thalassaemia as the result of a major deletion on its 5′ side suggests that the expression of genes in the ε-γ-δ-β cluster is influenced through some form of interaction in *cis*[11,21]. The sequential arrangement of the ε, γ, δ and β genes parallels ontogenesis, and there is recent evidence that γ-globin synthesis precedes δ- and β-globin synthesis in the maturing erythrocyte[22,23]. This suggests the possibility that these genes are sequentially activated and inactivated by a mechanism, such as a progressive change in chromatin structure, acting only within the chromosome. It would then be expected that at least some of the genetic lesions affecting globin gene activity would be polar in nature, exerting their effect by blocking the progress of gene activation. Thus the common feature in both kinds of $^G\gamma$-$(\delta\beta)^\circ$-thalassaemia would be that the deletions start in a similar place, and the precise form of genetic rearrangement downstream from this site would be largely irrelevant. This idea could also be applied to the non-deletion forms of HPFH in which varying proportions of γ- and β-globin chains are present in the mature adult erythrocyte. Here genetic lesions undetectable by restriction enzyme analysis could, because of their position in the ε-γ-δ-β globin cluster, delay the progress of gene activation during erythrocyte maturation so that the mature erythrocyte would contain a higher than normal proportion of γ-globin.

We thank A. B. Amin for providing the $^G\gamma$-$(\delta\beta)^\circ$-thalassaemia blood samples, T. Maniatis for the clone HγG4 and the *Pst* β and *Pst* δ subclones, B. Forget for the plasmids JW151 and JW102, R. A. Flavell, J. Edwards and D. Roberts for helpful discussion, O. Smithies and B. Forget for sending us their results before publication, S. Goodbourn for preparing the *Pst* β and *Pst* δ subclones, and H. Ayyub for technical assistance.

Received 1 December 1980; accepted 9 February 1981.

1. Proudfoot, N. J. & Baralle, F. E. *Proc. natn. Acad. Sci. U.S.A.* **76,** 5435–5439 (1979).
2. Flavell, R. A., Kooter, J. M., DeBoer, E., Little, P. F. R. & Williamson, R. *Cell* **15,** 25–41 (1978).
3. Little, P. F. R., Flavell, R. A., Kooter, J. M., Annison, G. & Williamson, R. *Nature* **278,** 227–231 (1979).
4. Fritsch, E. F., Lawn, R. M. & Maniatis, T. *Nature* **279,** 598–603 (1979).
5. Fritsch, E. F., Lawn, R. M. & Maniatis, T. *Cell* **19,** 959–972 (1980).
6. Jeffreys, A. J., Craig, I. W. & Francke, V. *Nature* **281,** 606–608 (1979).
7. Amin, A. B. *et al. Br. J. Haemat.* **43,** 537–548 (1979).
8. Ottolenghi, S. O. *et al. Nature* **278,** 654–656 (1979).
9. Bernards, R., Kooter, J. M. & Flavell, R. A. *Gene* **6,** 265–280 (1979).
10. Orkin, S. H., Alter, B. P. & Altay, C. *J. clin. Invest.* **64,** 866–869 (1979).
11. Bernards, R. & Flavell, R. A. *Nucleic Acids Res.* **8,** 1521–1534 (1980).
12. Tuan, D., Murnane, M. J., de Riel, J. K. & Forget, B. G. *Nature* **285,** 335–337 (1980).
13. Smithies, O. *Cold Spring Harb. Symp. quant. Biol.* **29,** 309–319 (1964).
14. Chow, L., Kahmann, R. & Kamp, D. *J. molec. Biol.* **113,** 591–609 (1977).
15. Roeder, G. S. & Fink, G. R. *Cell* **21,** 239–240 (1980).
16. Slightom, J. L., Blechl, A. E. & Smithies, O. *Cell* **21,** 627–638 (1980).
17. Lawn, R. M., Efstratiadis, A., O'Connell, C. & Maniatis, T. *Cell* **21,** 647–651 (1980).
18. Spritz, R. A., DeRiel, J. K., Forget, B. G. & Weissman, S. M. *Cell* **21,** 639–646 (1980).
19. Weatherall, D. J. & Clegg, J. B. *The Thalassaemia Syndromes* 3rd edn (Blackwell Scientific, Oxford, 1981).
20. Huisman, T. H. J. *et al Ann. N.Y. Acad. Sci.* **232,** 107–124 (1974).
21. van der Ploeg, L. H. T. *et al. Nature* **283,** 637–642 (1980).
22. Papayannopoulou, T., Kalmantis, T. & Stamatoyannopoulos, G. *Proc. natn. Acad. Sci. U.S.A.* **76,** 6420–6424 (1979).
23. Comi, P. *et al. Proc. natn. Acad. Sci. U.S.A.* **77,** 362–365 (1980).
24. Jeffreys, A. J. *Cell* **18,** 1–10 (1979).
25. van der Ploeg, L. H. T. & Flavell, R. A. *Cell* **19,** 947–958 (1980).
26. Orkin, S. H., Old, J. M., Weatherall, D. J. & Nathan, D. G. *Proc. natn. Acad. Sci. U.S.A.* **76,** 2400–2404 (1979).
27. Mears, J. G. *et al. Proc. natn. Acad. Sci. U.S.A.* **75,** 1222–1226 (1978).
28. Wahl, G. M., Stern, M. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **76,** 3683–3687 (1979).
29. Gross-Bellard, M., Oudet, P. & Chambon, P. *Eur. J. Biochem.* **36,** 21–38 (1973).
30. Southern, E. M. *J. molec. Biol.* **98,** 503–517 (1975).
31. Kan, Y. W. & Dozy, A. M. *Proc. natn. Acad. Sci. U.S.A.* **75,** 5631–5635 (1978).
32. Maniatis, T., Jeffrey, A. & Kleid, D. G. *Proc. natn. Acad. Sci. U.S.A.* **72,** 1184–1188 (1975).
33. Orkin, S. H. *et al. Cell* **17,** 33–42 (1979).
34. Laskey, R. A. & Mills, A. D. *FEBS Lett.* **82,** 314–316 (1977).
35. Wilson, J. T. *et al. Nucleic Acids Res.* **5,** 563–581 (1978).

LETTERS

Eternity matters

Don N. Page & M. Randall McKee

Departments of Physics and Astronomy, The Pennsylvania State University, University Park, Pennsylvania 16802, USA

The matter content of an ever-expanding Friedmann universe will never become negligible compared with the radiation, assuming electrons and positrons ($e^\pm$s) are absolutely stable, even if all other massive particles are unstable. The possible annihilation of $e^\pm$s in a $k = 0$ model occurs so slowly that the radiation density can never greatly exceed the matter density. For a very homogeneous distribution of $e^\pm$s which cool at just the expansion rate, the asymptotic ratio of matter to radiation density in a $k = 0$ universe is calculated here to be $(\sqrt{105}-3)/12 = 0.6039$. Heating by collisional recombination would probably keep the matter density far higher. A certain amount of clumping might make the annihilation occur faster, but larger inhomogeneities are likely to lead to black holes which could be even more effective in keeping the universe from becoming radiation dominated.

Several hypotheses[1–4] have been proposed to describe the future of an ever-expanding universe. To make some definite predictions, we assume here an open universe that is approximately homogeneous and isotropic (described by a Friedmann model) and which obeys Einstein's equations with no cosmological constant. (Without some such assumptions we could not rule out any conceivable influence coming in from beyond our present particle horizon.) We also assume that charge is absolutely conserved (so $e^\pm$s, as the lightest charged particles, are stable) and that no phase transitions bring in fundamentally new physics at arbitrarily low temperatures. Within such a model, the older picture of stray elementary particles, atoms, stars and black holes spreading out in space has been greatly altered by predictions of the decay of baryons[5–8] and black holes[9]. Thus it seems that the only stable particles may be electrons, positrons, neutrinos, photons and gravitons. In a $k = -1$ Friedmann model these particles would spread further and further apart and have a negligible chance of ever meeting after a sufficiently late time, so such a universe would always remain matters dominated. It has been argued[3] that a $k = 0$ universe would become radiation dominated after the thermal kinetic energy of the electrons and positrons ($e^\pm$s) drops below the Coulomb potential energy between pairs so that they become electrically bound and eventually annihilate. However, during this slow annihilation the energy in radiation is being redshifted, so a more careful analysis is needed to compare the decreasing mass density of matter with that of radiation.

Assume a $k = 0$ Friedmann universe expands according to $R \propto t^\alpha$ where $1/2 \leq \alpha \leq 2/3$, each extreme representing complete domination by either radiation or matter respectively. The direct collisional annihilation of free $e^\pm$s has a low-velocity cross-section[10] $\sigma \propto v^{-2} \propto t^{2\alpha}$ and hence a probability of occurring which becomes negligible at late times if $\alpha > 1/2$. It is more probable for $e^\pm$s to bind into positronium first. However, two-body radiative 'recombination' has a total cross-section[11] only logarithmically greater than the direct collisional annihilation and is, therefore, also negligible. Collisions of neutral particles are even more rare, so any stable massive neutral particles would keep the universe matter dominated. To continue the argument against radiation domination, we set this case aside and assume $e^\pm$s are the only stable massive particles.

Three-body collisional recombination of positronium is the dominant mechanism leading to $e^\pm$ annihilation at late times. Let $N = xt^{-3\alpha}$ be the number density of all $e^\pm$s that have not

0028-0836/81/190044—02$01.00

annihilated and $N_e = yt^{-3\alpha}$ be that of free e$^\pm$s. Then if the temperature of free e$^\pm$s drops as $T \propto t^{-2\alpha}$,

$$\begin{aligned} t^{-3\alpha}\ dy/dt &= -N_e^3 \sum_n C_{i,n} \\ &\sim -y^3 t^{-9\alpha}(m_e^{-4} e^{-4} T^{-1} n_{\max}^7) \\ &\propto -y^{11/6} t^{-7\alpha/2} \end{aligned} \tag{1}$$

using the collisional ionization cross-section[12,13] and detailed balance to get the recombination rate coefficient $C_{i,n}$ into the nth level and then summing over all levels up to $n_{\max} \sim m_e^{1/2} e y^{-1/6} t^{\alpha/2}$ where the size of the positronium atom equals the separation between free e$^\pm$s. (There is no problem forming new positronium atoms surrounding smaller ones, because as neutral objects the smaller atoms will have a negligible effect on the larger ones.) Thus at late times the fraction of the original e$^\pm$s that are still free is proportional to $y \propto t^{-3(2-\alpha)/5} \propto n_{\max}^{-6\beta}(t)$, where

$$\beta = (2-\alpha)/(2+4\alpha) \tag{2}$$

After forming in levels around $n_{\max}(t)$, the positronium spirals into the ground state by radiative transitions, annihilating at a time $t_a \propto n_{\max}^6(t) \propto y^{-1/\beta}(t)$. Hence the fraction of e$^\pm$s that have not annihilated by time t_a is proportional to $x(t_a) \approx y(t) \propto t_a^{-\beta}$. This decreases more slowly than the redshifting black-body radiation energy, so the e$^\pm$ mass density ρ_e always dominates that radiation density. However, the e$^\pm$ annihilation gives a coupling to the total radiation density ρ_r by

$$d\rho_e/dt = (-3H - \beta/t)\rho_e, \qquad d\rho_r/dt = -4H\rho_r + (\beta/t)\rho_e \tag{3}$$

In a $k = 0$ universe dominated by these two densities so that $H^2 \equiv \dot{R}^2/R^2 = 8\pi(\rho_e + \rho_r)/3$, the asymptotic solution of equation (3) is

$$\begin{aligned} &\rho_e = (1-2\beta)(1-\tfrac{1}{2}\beta)/6\pi t^2, \\ &\rho_r = \tfrac{3}{2}\beta(1-\tfrac{1}{2}\beta)/6\pi t^2, \qquad R \propto t^{(2-\beta)/3} \end{aligned} \tag{4}$$

This gives the $R \propto t^\alpha$ expansion assumed above with equations (2) and (4) implying

$$\alpha = (3+\sqrt{105})/24 = 0.5520, \qquad \beta = (13-\sqrt{105})/8 = 0.3441 \tag{5}$$

Then $\rho_e/\rho_r = (2-4\beta)/3\beta = (\sqrt{105}-3)/12 = 0.6039$. Even though the e$^\pm$s are annihilating, their density is thus never completely dominated by the radiation density.

The annihilation of e$^\pm$s may be even slower than this, because the collisional recombination of positronium will probably keep the temperature of the free e$^\pm$s from dropping as fast as $t^{-2\alpha}$. In fact, if the temperature is kept comparable with the energy that is released when each e$^\pm$ pair becomes bound, $T \sim m\, e^4 n_{\max}^{-2}$, then equation (1) becomes $t^{-3\alpha}\ dy/dt \propto -y^{3/2} t^{-9\alpha/2}$. This gives $\beta = 1 - 3\alpha/2$ with the asymptotic solution (4) implying $\alpha = 2/3$, $\beta = 0$. In this case the number of e$^\pm$s decreases only logarithmically with time, and the universe stays completely matter dominated.

All this assumes a nearly perfect Friedmann model with negligible inhomogeneities. Clumping of the e$^\pm$s will make their annihilation occur faster but the annihilation itself will tend to destroy the inhomogeneities. The asymptotic behaviour will depend on the spectrum of the inhomogeneities at large scales and their growth compared with their destruction. Inhomogeneities with vorticity insufficient to prevent them from collapsing gravitationally may form black holes if they are large enough to collapse completely before their e$^\pm$s annihilate. (This requires a mass $M \gtrsim e^4 m_e^{-3} \sim 10^{25} M_\odot$.) If black holes (which are here classified as matter as their rest mass dominates their energy) form with a spectrum[14] $\propto M^{-\gamma}\, dM$ up to arbitrarily high mass M, their decay[9] with lifetime $\propto M^3$ will produce a coupling to radiation with $\beta_H = (\gamma - 2)/3$. (We still assume an approximate $k = 0$ Friedmann model on sufficiently large scales.) If $\beta_H > \beta$, the black hole density ρ_H will decrease faster than ρ_e and equation (4) will apply. If not, ρ_H will equal or dominate ρ_e and equation (4) will have ρ_e replaced by ρ_H and β replaced by β_H. However, the black holes will probably coalesce into ever bigger and bigger holes. Indeed, even the radiative 'recombination' cross-section $\sigma \sim M^2 v^{-18/7}$ for two black holes to become bound and eventually spiral together will cause each black hole to undergo an infinite number of coalescences (if it gets big enough not to evaporate between collisions) for $\alpha \leqslant 0.7$, unlike the case for e$^\pm$ radiative recombinations which require $\alpha \leqslant 0.5$ to be effective. Multibody interactions are likely to be even more effective, as gravity is always attractive. Thus black holes may keep a $k = 0$ universe always entirely matter dominated.

If, on the other hand, there is a maximum mass to the black holes that form, then they will all evaporate, producing mostly massless radiation. This could make the universe radiation dominated temporarily, but eventually the stray matter or the roughly 10^{44} e$^\pm$s emitted during the final stages of each black hole decay[15] will become comparable with the radiation density. Thus in any case radiation will never completely dominate the density of an ever-expanding approximately Friedmann universe; matter will always be important.

More details of the e$^\pm$ annihilation in a $k = 0$ universe, including the spectrum of the enormous number of photons produced, will be given elsewhere[16].

D.N.P. appreciates support from NSF grant PHY79-18430.

Received 12 December 1980; accepted 3 March 1981.

1. Davies, P. C. W. *Mon. Not. R. astr. Soc.* **161,** 1–5 (1973).
2. Islam, J. N. *Q. Jl R. astr. Soc.* **18,** 3–8 (1977); *Sky Telescope* **57,** 13–18 (1979).
3. Barrow, J. D. & Tipler, F. J. *Nature* **276,** 453–459 (1978).
4. Dyson, F. J. *Rev. mod. Phys.* **51,** 447–460 (1979).
5. Pati, J. C. & Salam, A. *Phys. Rev.* **D8,** 1240–1251 (1973); *Phys. Rev. Lett.* **31,** 661–664 (1973).
6. Georgi, H. & Glashow, S. L. *Phys. Rev. Lett.* **32,** 438–441 (1974).
7. Georgi, H., Quinn, H. R. & Weinberg, S. *Phys. Rev. Lett.* **33,** 451–454 (1974).
8. Zel'dovich, Ya. B. *Phys. Lett.* **59A,** 254 (1976); *Soviet Phys. JETP* **45,** 9–11 (1977).
9. Hawking, S. W. *Nature* **248,** 30–31 (1974); *Commun. Math. Phys.* **43,** 199–220 (1975).
10. Bransden, B. H. *Case Studies in Atomic Collision* I (eds McDaniel, E. W. & McDowell, M. R. C.) (North-Holland, Amsterdam, 1969).
11. Kramers, H. A. *Phil. Mag.* **46,** 836–871 (1923).
12. Thomson, J. J. *Phil. Mag.* **23,** 449–457 (1912).
13. Percival, I. C. & Richards, D. *Adv. atom. molec. Phys.* **11,** 1–82 (1975).
14. Carr, B. J. *Astrophys. J.* **201,** 1–19 (1975).
15. Page, D. N. *Phys. Rev.* **D16,** 2402–2411 (1977).
16. Page, D. N. & McKee, M. R. Preprint (Pennsylvania State University, 1980).

Ultra-heavy cosmic ray studies with Ariel VI

P. H. Fowler, R. N. F. Walker, M. R. W. Masheder, R. T. Moses & A. Worley

H. H. Wills Physics Laboratory, University of Bristol,
Tyndall Avenue, Bristol BS8 1TL, UK

The satellite Ariel VI was launched on 3 June 1979 into a near-circular 625 km 55° inclination orbit. The Bristol experiment on board is of a novel design using a spherical gas scintillation chamber with an integral plastic Cerenkov element[1]. The data presented here were collected during an exposure of ~360 m^2 sr days with essentially 100% recording efficiency for events with $Z \geqslant 38$. The abundance spectrum, particularly for $Z \geqslant 50$, suggests that the cosmic ray source is significantly enriched in r-process material relative to the Solar System.

Charge estimates for ultraheavy nuclei ($Z \geqslant 30$) are made on the assumption that both components of the signal increase as Z^2 and use a daily calibration at $_{26}$Fe based on about 15,000 events. Both components of the signal have remained essentially

Table 1 Comparison of the spectrum $N(Z_{est})$ from the present experiment with those of previous experiments on balloons and Skylab

Charge	Present experiment		Balloon data (top of atmosphere)		Skylab	
26	10^6		10^6		10^6	
50	4.3 ± 1.4		*7.5 ± 2.3			
52	6.5 ± 1.5		*6.2 ± 2.0			
54	2.2 ± 1.1		*5.5 ± 1.7			
56	7.7 ± 1.5		*5.4 ± 1.7			
58	1.9 ± 1.1	4.1 ± 1.3	*5.8 ± 1.7	10.6 ± 2.1		
60	2.2 ± 1.0		*4.8 ± 1.4			
62	1.8 ± 0.8	7.8 ± 1.3	*3.1 ± 1.0	6.1 ± 1.2		
64	3.1 ± 0.9		*1.8 ± 0.7		*3.0 ± 1.0	
66	2.9 ± 1.0		*1.2 ± 0.5		*2.7 ± 0.9	
68	0.1 ± 0.6	2.9 ± 0.9	0.6 ± 0.4	2.4 ± 0.6	*0.8 ± 0.4	3.3 ± 0.8
70	1.5 ± 0.7		0.8 ± 0.3		*1.6 ± 0.5	
72	1.3 ± 0.7		1.0 ± 0.3		*0.9 ± 0.4	
74	1.8 ± 0.8	5.8 ± 1.1	1.1 ± 0.3	4.4 ± 0.8	0.9 ± 0.3	3.8 ± 0.6
76	0.9 ± 0.6		1.5 ± 0.4		1.6 ± 0.4	
78	3.1 ± 0.9		1.8 ± 0.5		1.3 ± 0.4	
80	0.9 ± 0.7	3.0 ± 0.9	1.3 ± 0.4	3.0 ± 0.6	0.9 ± 0.3	2.3 ± 0.5
82	1.7 ± 0.8		0.9 ± 0.4		0.6 ± 0.3	
84	0.4 ± 0.6		0.8 ± 0.3		0.8 ± 0.3	
86	—	0.5 ± 0.3	0.5 ± 0.3	1.9 ± 0.5	0.2 ± 0.2	0.8 ± 0.25
88	0.2 ± 0.2		0.3 ± 0.2		0.0	
≥90	0.3 ± 0.2		1.1 ± 0.4		0.6 ± 0.3	
Based on	137 events $50 < Z < 65$		84 events $50 < Z < 65$		104 events $Z > 65$	
and	102 events $Z > 65$		96 events $Z > 65$			

The balloon data have been corrected to yield the fluxes at the top of the atmosphere for a better comparison with the space data. The nominal exposures involved for the present experiment and the balloon data are each ~400 m^2 sr days, and for Skylab ~700 m^2 sr days. The balloon data is all obtained at high geomagnetic latitude and under ~4 g cm^{-2} of overlying air. Numbers using wider charge bins are also given for high Z because of the correlated errors in our data.

* The value is subject to significant and rapidly charge-dependent corrections due to threshold effects for the tracks in plastic.

constant since launch. The linearity has been confirmed for the lower charges with the observation of peaks for all the even charges $8 \leq Z \leq 26$ and the charge resolution at iron is $\sigma \sim 0.4$ charge. Data collected in the neighbourhood of the South Atlantic Anomaly and during occasional periods of interference from high particle rates in the auroral belts have been rejected.

Figure 1 shows the charge spectrum of apparent charge Z_{app} near $_{26}$Fe for events collected in regions where the geomagnetic cutoff is between 3 and 5 GV. The tail above $Z = 26$, caused by the relativistic rise in the scintillation signal, reduces the $_{28}$Ni peak to a shoulder and smothers the contribution from $_{30}$Zn. This process results in an uncertainty in charge ΔZ which is proportional to Z and is the dominant term at high Z. Nuclear fragmentation within the thin (~1 g cm^{-2}) detector will produce a tail to low Z_{app} which is masked by the form of the charge spectrum. Additionally, both electron showers generated in the spacecraft body and stopping iron nuclei have the ability to mimic high charge events, but these spurious events can be eliminated by suitable selection of the data. A plastic scintillator guard counter intercepts all electron shower events from the spacecraft and, unavoidably, ~30% of genuine events; all these events must be rejected to remove the showers. Electron shower contamination was found to be sufficiently low for $Z \geq 62$ that all events were accepted regardless of the guard counter signal. Spurious apparently high charge events due to stopping iron group nuclei can be produced by the very small fraction of such nuclei whose entry ranges are both ≳7 g cm^{-2} and exceed by only ~0.1 g cm^{-2} the total path in the detector, such that an excessively large energy deposition occurs during their passage through the final 0.01 g cm^{-2} of gas. The extent of the excess energy loss does, however, have a well-defined upper limit, and only events with $Z_{app} \leq 50$ may be produced in this way by $_{26}$Fe or $_{28}$Ni. For this reason the charge spectrum has been divided at $Z_{app} = 50$ and high latitude events excluded from the lower charge accumulation.

The spectrum of apparent charge that results from the experiment contains the effects of the relativistic rise. These effects can be removed statistically to give a spectrum of estimated charge Z_{est}. It is assumed that the spectra of energy per nucleon for ultraheavy cosmic rays and for $_{26}$Fe nuclei are indistinguishable[2]. Further evidence for the similarity of the energy spectra is obtained in this experiment from comparison of the flux values as a function of the geomagnetic cutoff. Table 1 compares our spectrum of Z_{est} with results from the Skylab experiment[3] and balloon work[2]. The comparison is valid only for $Z \geq 50$ due to the threshold effects of track registration in the Lexan plastic detectors. There are features common to the three spectra, the most significant being the overall flux values, the prominent peak at $Z \sim 78$ and a comparable abundance of nuclei with $58 \leq Z \leq 66$. The Skylab estimate for $Z \sim 66$ was considered by its authors to be uncertain due to the proximity of the threshold referred to above. The general agreement between the sets of results supports the charge scale used in the Lexan track analysis in both the Skylab and balloon experiments.

Table 2 presents the data from this experiment for the whole available charge range. The spectra of Z_{app} and Z_{est} are given together with the original number of events on which each entry was based. Also shown is our estimate of the charge spectrum of secondaries produced by interactions in the interstellar medium that are included among the nuclei detected. The difference between the two figures is the estimated number of residual primaries. This is extrapolated back to the source region to give our estimated source spectrum which again is normalized to Fe = 10^6 and which can be directly compared with the Solar System abundances compiled by Cameron[4]. We used the simplest form of 'leaky box' model for cosmic ray propagation which ignores ionization energy loss and assumes all cross-sections for interaction with interstellar hydrogen to be independent of energy. Cross-sections for nuclear interactions were based on those given for 2.3 GeV per nucleon by R. Silberberg (personal communication). The escape mean free path was taken as 5.0 g cm^{-2} of hydrogen. Clearly, except for the highest charges with $Z \geq 78$, a substantial fraction of the observed ultraheavy nuclei are secondaries arising from interactions in interstellar space. This will be the case whenever the amount of matter traversed by the detected particles is comparable with their interaction mean free paths.

Several authors have concluded that the ultraheavy cosmic ray spectrum seems to be enriched in material which has undergone nucleosynthesis by the rapid neutron capture process. More recently it has been suggested[5] that at least for the elements $Z \leq 28$ the cosmic ray source spectrum could be simply explained as a sample of normal interstellar material, modified only by the effects of differing ionization potentials for the various elements. The cosmic ray source spectrum in the ultraheavy domain would on this assumption have much in common

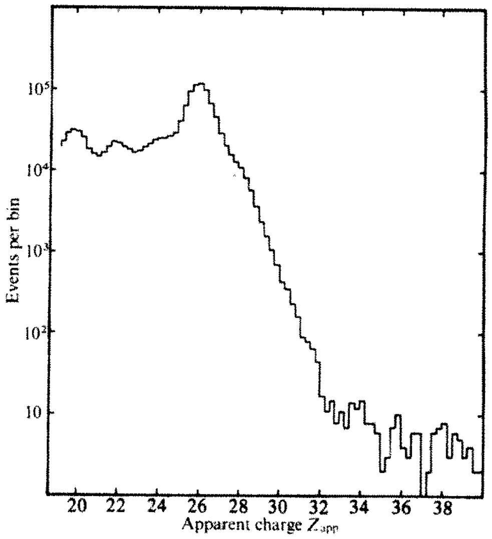


Fig. 1 Histogram of apparent charge of events collected during 110 days of operation for cutoffs between 3GV and 5GV. The sample contains $\sim 7 \times 10^5$ iron nuclei.

Table 2 Estimates of the spectra at several stages between detection and the computed source composition together with the Solar System composition of Cameron

Charge	Present experiment					Solar System abundances
	$N(Z_{app})$	$N(Z_{est})$	N_{sec}	N_{source}		
Fe = 26	6.3×10^6	6.3×10^6	6.3×10^6	10^6		10^6
34	557 (213)	526	217	59±9		83
36	327 (125)	290	145	29±7		54
38	298 (114)	301	97	41±7		31
40	196 (74)	175	58	24±6		17
42	93 (35)	71	43	6±4		4.9
44	66 (25)	63	37	5.7±3.8	7.8±4.2	2.3
46	42 (16)	36	27	2.1±3.1		1.9
48	48 (18)	51	26	6.0±3.5	4.8±3.4	2.1
50	32 (26)	27	32	−1.2±2.1		4.4
52	38 (31)	41	19.6	5.3±2.3	11.1±3.3	8.1
54	20 (16)	13.8	20	−1.7±1.8		7.4
56	40 (33)	48	20	7.5±2.5		5.8
58	20 (16)	12.0	6.6	1.5±1.9	2.6±2.2	1.6
60	15 (12)	14.0	10.0	1.1±1.7		0.98
62	12	11.1	13.3	−0.6±1.4	4.9±2.4	0.32
64	17	19.4	12.5	2.0±1.7		0.56
66	18	18.5	6.8	3.5±1.8		0.50
68	6	0.8	6.0	−1.6±1.1	−0.1±1.3	0.33
70	8	9.3	5.9	1.1±1.3		0.26
72	8	8.1	6.7	0.4±1.3		0.22
74	10	11.1	3.9	2.3±1.5	7.6±2.3	0.37
76	7	5.6	5.4	0.1±1.3		1.20
78	15	19.3	3.4	5.2±2.0		2.08
80	9	6.0	2.5	1.2±1.6	5.3±1.9	0.46
82	10	10.7	0.9	3.4±1.7		3.07
84	5	2.4	0.3	0.7±1.2		0.08
86	3	—	—	—		—
88	1	1	—	0.3±0.3		—
⩾90	2	2	—	0.7±0.5		0.08

For the highest charges all the detected events were used. For $Z \leqslant 60$, where selections were necessary, the values in brackets refer to the number of events utilized. The errors shown are due to Poisson statistics alone and are quite strongly negatively correlated in adjacent charge bins. Wider bins are therefore used at higher charges to display the features of the charge spectrum.

with the Solar System composition at all charges and, indeed, the overall abundance in our source spectrum is very similar to the Solar System value, a point which has been apparent from previous experiments for several years. In addition, the main features of the Solar System composition, namely the peaks at $Z \sim 54$ and $Z \sim 80$, are both clearly present in our data. However, the abundance ratio between $52 \leqslant Z \leqslant 56$ and $74 \leqslant Z \leqslant 84$ is 0.86±0.30 compared with a Solar System value of 2.9, and the predominant species in the upper peak is $_{78}$Pt rather than $_{82}$Pb. A striking anomaly also occurs at $Z \sim 64$ where the cosmic rays seem to be ~4 times over-abundant. These discrepancies cannot be attributed to the effects of ionization potentials. If as an alternative we suppose that much shorter mean free paths were to apply in the leaky box, the estimated source spectrum would approach the detected spectrum thereby reducing the excess for $Z \geqslant 74$ in the cosmic ray source. At the same time, however, the rare earths as a whole ($58 \leqslant Z \leqslant 72$) would become severely overabundant and the overall similarity to the Solar System abundances would not be improved.

There is a small Solar System abundance anomaly at $Z \sim 64 (A \sim 164)$ which has been attributed to fission fragments. Steinberg and Wilkins[6] consider that the feature may be produced by the heavier fragment resulting from asymmetric fission of superheavy nuclei with $A \sim 300$. Such nuclei are produced in nucleosynthesis models involving a cyclic r-process, but will be accompanied by substantially greater quantities of $A \sim 280$ nuclei. Our results seem inconsistent with this interpretation as our source spectrum contains neither a feature at $A \sim 280$ nor the excessive abundance at $A \sim 134 (Z \sim 54)$ which would result from the symmetric fission of the $A \sim 280$ nuclei advocated by Steinberg and Wilkins. We propose as an alternative that the $A \sim 280$ feature itself is responsible, by asymmetric fission, for the anomaly at $Z \sim 64$ in both the cosmic ray and Solar System abundances, the original mass being distributed between two peaks at $A = 105$ and $A = 164$ and perhaps 10 neutrons. Our data have as yet insufficient statistical weight to determine whether the region of the lighter peak, at $Z \sim 44$, is appropriately enhanced.

In the charge region $Z \geqslant 90$ we would have expected, on the basis of the Skylab and Solar System abundances, to have recorded respectively ~4 and ~0.2 such events in the present sample. Of the two events observed, the event with $Z = 99$ could be a perfectly normal actinide. Its light was well distributed among the photomultipliers of the detector sphere and the absence of a plastic scintillator signal rules out a shower origin in the spacecraft body. Because the highest charge event, with $Z = 114$, did activate the guard counter its interpretation as an electron shower cannot be excluded. Our conclusion that the cosmic ray source seems to be significantly enriched in the heavier r-process nuclei therefore rests on our discussion of the charge region $52 \leqslant Z \leqslant 84$.

Received 19 February; accepted 23 March 1981.

1. Fowler, P. H. *et al. Proc. 16th Int. Cosmic Ray Conf.* **12,** 338 (1979).
2. Fowler, P. H. *et al. Proc. 15th Int. Cosmic Ray Conf.* **11,** 165 (1977).
3. Shirk, E. K. & Price, P. B. *Astrophys. J.* **220,** 719 (1978).
4. Cameron, A. G. W. Preprint no. 1357 (Harvard-Smithsonian Center for Astrophysics, 1980).
5. Cassé, P. & Goret, P. *Astrophys. J.* **221,** 703 (1978).
6. Steinberg, E. P. & Wilkins, B. D. *Astrophys. J.* **223,** 1000 (1978).

Albedo change by man: test of climatic effects

Gerald L. Potter, Hugh W. Ellsaesser, Michael C. MacCracken & James S. Ellis

Lawrence Livermore National Laboratory, University of California,
PO Box 808 Livermore, California 94550, USA

Sagan *et al.*[1] have suggested that the cumulative impact of anthropogenic albedo changes may have contributed to global climate changes in the past and that its effect may be continuing. Using a statistical dynamic climate model (SDM) with more realistic surface albedo changes than previously used[2], we have computed the combined impact of desertification of the Sahara and deforestation of the tropical rain forest. While the model computed a surface cooling of 0.6 K for the Northern Hemisphere, the global mean of ~0.2 K was substantially less than the 1 K suggested by Sagan *et al.*[1]. We infer that man's cumulative impact on planetary surface albedo over the past few thousand years has had a small and probably undetectable effect on global climate.

Earlier numerical model experiments[2–4], suggesting that the main effect of man's past activities should have been a general increase of surface albedo resulting in lower surface temperatures, were not designed to incorporate more than single impacts, such as tropical deforestation or desertification. The suggestion of Sagan *et al.*[1] that combined effects may have long term global ramifications seemed sufficiently serious to warrant investigation in a more sophisticated model. We have therefore used the Lawrence Livermore National Laboratory (LLNL) two-dimensional SDM to calculate the combined climatic effect of desertification in the Northern Hemisphere subtropics and deforestation of the tropical rain forest. Other surface modifications suggested by Sagan *et al.*[1] were not included as they involved relatively small surface areas and there is less agreement as to their potential impact on climate.

The LLNL two-dimensional SDM model[5] is similar in many ways to general circulation models using the basic conservation equations to calculate zonal mean fields of pressure, temperature, wind and humidity. The model has nine pressure layers and a 10°-latitude grid. The surface area of each latitude zone is divided into land (of differing types and elevations) and ocean (partially covered by ice). Sea ice (with the appropriate albedo) is formed when the ocean temperature drops below 271.4 K. New ice spreads at a minimum thickness of 20 cm until it

0028-0836/81/190047—03$01.00

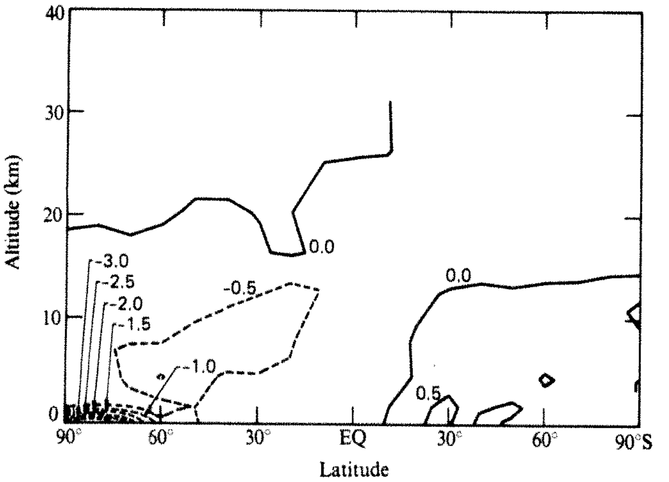


Fig. 1 Latitude–height temperature difference (perturbed minus control). Contour interval = 0.5 K.

Table 1 Model albedo changes from Sagan *et al.*[1]

Process	Land type change	Area [km^2] (% band)	Centred on latitude bands	Surface albedo change
Desertification	savanna to desert	9×10^6 (22%)	20° N	0.16→0.35
Tropical deforestation	forest to field savanna	7×10^6 (5% of 10° S, 10% of Equator)	10° S and Equator	0.07→0.16

occupies the maximum allowable fractional coverage (leaving some open ocean to simulate open leads). If the energy deficit continues the sea ice then increases in depth. Conversely, an energy excess would melt the ice before the water warms. The melt water (with the appropriate albedo) can refreeze on top of the older ice and thus increase the ice depth. Fractional grid zone assignments of various surface types with corresponding physical characteristics (for example, albedo, heat capacity, bulk transfer coefficients) enable the sensitivity of the model to modifications of these parameters to be tested.

Surface fluxes of sensible and latent heat are calculated using the vertical gradients of water vapour and temperature, the surface wind velocity and low-level convective activity. Clouds are randomly correlated vertically and are a prescribed latitude and height dependent function of relative humidity[6] (for further details see refs 4 and 7).

The sensitivity, or response to an imposed perturbation, of the model is conveniently measured by its response to an increase of its solar constant. The usual unit of comparison is the sensitivity parameter β (equal to 100 times the temperature change resulting from a 1% increase in solar flux), and experiments with our SDM give a global sensitivity factor $\beta = 70$ K. This is less than half the value quoted for simple energy balance models[8,9] and the Geophysical Fluid Dynamics Laboratory sector model[10]. At least some of this reduced sensitivity results from use of a global model and from the greatly reduced sensitivity of the Southern Hemisphere, which exhibits significantly less ice-albedo and land-surface albedo feedback.

The non-seasonal mode of the LLNL SDM simulates the annual average climate by using a fixed solar zenith angle for each latitude calculated to supply the appropriate annual average solar flux. By eliminating the diurnal and seasonal fluctuations, the model approaches a pseudo-equilibrium state after ~200 days of model integration time. This reflects the adjustment time implied by the heat capacities of the atmosphere, land, snow, ice and arbitrarily thin oceanic mixed layer. For this experiment, two model equilibrium states were achieved: a control case with standard surface conditions corresponding to a pristine environment; a perturbed case corresponding to the suggestion of Sagan *et al.*[1] in terms of tropical deforestation and desert denudation (see Table 1). The surface bulk transfer coefficients, the thermal and hydrologic characteristics were not altered; the only surface parameter modified for those areas perturbed was surface albedo which, after modification, was held constant. This reduced the degrees of freedom for the SDM.

The response to this change in albedo was generally consistent with previous surface albedo modification sensitivity studies[2,3]. Figure 1 shows the temperature difference (perturbed minus control) between the two cases. Generally, the Northern Hemisphere troposphere cooled, with a maximum change due to ice-albedo feedback near the surface at 70° N.

There seem to be two important factors in the chain of events leading from surface modification to cooling in the Northern Hemisphere. The first is the direct reduction in surface absorption in the zones of increased surface albedo (see Table 1). The change at 20° N was sufficiently strong to strengthen the temperature gradient between the Sahara[3] and the Equator which caused the northern Hadley cell to strengthen and shift southwards. This in turn weakened the hydrological activity at 20° N with less precipitable water associated convection and latent heat release, accompanied by maximum cooling in the upper troposphere, lowering of the tropopause and warming of the lower stratosphere. With less solar energy absorbed in the subtropics there was reduced poleward transport of sensible and latent heat and a general cooling of the higher northern latitudes. The second factor cooling the Northern Hemisphere was the ice-albedo feedback due to increased sea-ice coverage, primarily at 70° N.

Generally, the cloud cover increased in the Northern Hemisphere where the surface temperature had cooled 0.6 K. This agrees with reports[11–13] of global increases in cloud amount with decreased surface temperatures (Table 2). One exception was a general cloud decrease at 20° N due to increased Hadley subsidence. This latitude received the largest surface-albedo perturbation and hence the highest directly resultant surface temperature decrease[14].

In the Southern Hemisphere, only a small fraction of the latitude band at 10° S also had an albedo increase from deforestation. Figure 2 shows the zonal average conversion from potential to kinetic energy due to vertical motion. The heated air in the low level equatorial Hadley cell rises converting potential to kinetic energy. In the higher latitudes the sinking motion converts back to potential energy. This conversion term gives an

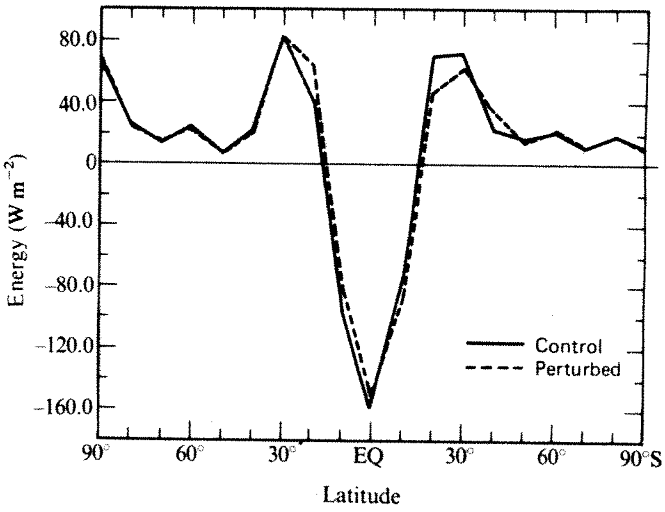


Fig. 2 Kinetic to potential energy conversion through vertical motion.

indication of the intensity of the Hadley cell. In the perturbed case, the Hadley cell shifted southwards with the ascending branch weakening slightly and the southern descending branch weakening more. The southward shift brought increased ascent to the southern subtropics. This with the weakened descending branch permitted increased convective activity, leading to clouds with more vertical structure and less horizontal extent. The total cloud coverage thus decreased at 30° S, allowing increased insolation. This increase, along with the albedo increase at the Equator and 10 °S, weakened the subtropical temperature gradient and presumably further weakened the Southern Hemisphere Hadley cell. However, other than the additional solar flux at 30° S, there seems little reason for the slight warming of the Southern Hemisphere.

To assess more fully the relative importance of the various model responses to the imposed increase in surface albedo, we can examine the contribution to the change in solar (S) and outgoing longwave (F) radiation at the top of the atmosphere as given by

$$\delta S = \delta S_r + \delta S_{A_c} + \delta S_{\alpha_s} + \delta S_f$$

$$\delta F = \delta F_T + \delta F_r + \delta F_{A_c} + \delta F_f$$

where r, T, A_c, α_s and f identify components due to precipitable water, temperature, cloud amount, surface albedo and fractional change in ice coverage, respectively. These individual contributions have been calculated by recomputing the planetary radiation balance keeping all of the input data fields fixed (control case) except for the perturbed data field of interest (for example, the temperature, water vapour, clouds or surface albedo from the modified albedo case).

Table 3 shows the change in solar, long-wave and net radiation due to the separated changes in various contributing factors; the two major components in changing the planetary energy balance are clouds and surface albedo. The mean global cloud fraction decreased 0.011, resulting in an increase in long-wave loss to space and an increase in solar radiation absorbed by the Earth–atmosphere system. The surface albedo increased as a result of both the imposed albedo modification and the ice-albedo feedback, decreasing the solar radiation absorbed by the previously clear portion of the Earth–atmosphere system. The mean surface temperature decrease reduced the long-wave loss to space because of the cooler radiating temperature.

Surprisingly, water vapour was the least important component in the long-wave response, although it did produce a reduction in long-wave loss to space. This small influence may be because water vapour experienced both vertical and horizontal redistributions in conjunction with the hemispheric asymmetries of the total model response. Thus the change in global mean precipitable water vapour was so small (−1.6%) that the radiative response was nearly insignificant.

As the model calculates sea ice, the ice cover fraction will differ from experiment to experiment. The effect of this change alone (all other effects held fixed) on the global energy balance is shown in Table 3*a*. Sea ice increased ~14% in the latitude band 65–75° N, which corresponds to an added sea-ice area of 2×10^6 km². The 0.26 W m^{-2} decrease in solar radiation represents the effect of sea ice alone, without cloud cover or surface albedo contributions. Long-wave radiation responded because of the larger area radiating at a colder sea-ice temperature (reducing the long-wave flux to space).

Table 2 Relationship between change in cloud cover and change in surface temperature in this experiment and those of Schneider *et al.*[11] (Northern Hemisphere only)

	δT	δA_c	$\partial A_c/\partial T$
LLNL SDM NH 4km cloud	−0.58	+0.004	−0.007
NCAR GCM 3 km cloud	−2.52	+0.015	−0.006

Table 3 *a* Global mean change in solar, outgoing long-wave, and net radiation (δS, δF, and δS–δF) in W m^{-2} of the control model due to insertion of only the perturbed field

	T	r	A_c	α_s	f
δS	—	0.01	1.25	−3.87	−0.26
δF	−0.53	−0.07	0.31	—	−0.03
$(\delta S - \delta F)$	0.53	0.08	0.94	−3.87	−0.23

b Global change in various parameters (perturbed minus control)

Parameter (ψ)	Change	$\Delta\psi/\psi$ (%)
Global surface temperature (T)	−0.2 K	
Mean precipitable water vapour (r)	−0.04 g cm^{-2}	(−1.6%)
Global cloud cover fraction (A_c)	−0.011	(−1.9%)
Surface albedo (α_s)	+0.018	(+8.8%)
Sea ice (f)	$+2 \times 10^6$ km²	(+13.5% in 67–75° N band

The sensitivity of the LLNL SDM to the combined effect of deforestation in the tropics and desertification in the subtropics reveals several interesting responses. The Hadley cell shifted southwards and the northern descending branch intensified. Conversely, the southern descending branch decreased in intensity, thereby reducing subsidence. This reduction in stability in the Southern Hemisphere produced clouds with more vertical and less horizontal extent, allowing more solar radiation to reach the Earth's surface and to increase temperatures at 30° S. The Northern Hemisphere temperature decreased because of the reduced solar radiation reaching the surface coupled with additional cooling due to ice-albedo feedback.

While the LLNL SDM computed a Northern Hemisphere surface cooling of 0.6 K, the global cooling of 0.2 K was substantially less than the 1 K suggested by Sagan *et al.*[1] for the cumulative effect of anthropogenic surface modifications over the pass several thousand years. Our model results suggest, therefore, that the global temperature response is significantly less and the Northern Hemisphere temperature response slightly less than that of the simple energy balance model cited by Sagan *et al.*[1].

We thank Dr W. J. Quirk for suggestions and comments, Connee S. Mitchell for helping with the calculations, Floy Weichel and Joni Schuld for typing the manuscript. This work was performed under the auspices of the US Department of Energy by the Lawrence Livermore National Laboratory under contract no. W-7405-Eng-48. This paper is a continuation of work presented at the International Workshop on Energy–Climate Interactions, Munster, West Germany in March 1980.

Received 15 September 1980; accepted 25 February 1981.

1. Sagan, C., Toon, O. B. & Pollack, J. B. *Science* **206,** 1363 (1979).
2. Charney, J. G., Stone, P. H. & Quirk, W. J. *Science* **187,** 434 (1975).
3. Ellsaesser, H. W., MacCracken, M. C., Potter, G. L. & Luther, F. M. *Q. Jl R. met. Soc.* **102,** 655 (1976).
4. Potter, G. L., Ellsaesser, H. W., MacCracken, M. C. & Luther, F. M. *Nature* **258,** 697 (1975).
5. MacCracken, M. C. thesis, Univ. California (1968).
6. London, J. *A Study of the Atmospheric Heat Balance. Final Rep.* Contract AF19(122)-165 (Research Division. College of Engineering, New York University, 1957).
7. Potter, G. L., Ellsaesser, H. W., MacCracken, M. C. & Luther, F. M. in *GARP Ser. Publ.* No. 22 (WMO, Geneva, 1979).
8. Cess, R. D. *J. atmos. Sci.* **33,** 1831 (1976).
9. Budyko, M. I. *Climate and Life* (Academic, New York, 1974).
10. Ramanathan, V. *J. atmos. Sci.* **34,** 1855 (1977).
11. Schneider, S. H., Washington, W. M. & Chervin, R. M. *J. atmos. Sci.* **35,** 2207 (1978).
12. Henderson-Sellers, A. *Nature* **279,** 786 (1979).
13. Roads, J. O. *J. atmos. Sci.* **35,** 753 (1978).
14. Bjerknes, J., Allison, L. J., Kreins, E. R., Godshall, F. A. & Warnecke, G. *Bull. Am. met. Soc.* **50,** 313 (1969).

Oxidative power of Mn(IV) and Fe(III) oxides with respect to As(III) in terrestrial and aquatic environments

D. W. Oscarson*, P. M. Huang*, C. Defosse† & A. Herbillon†

* Saskatchewan Institute of Pedology, University of Saskatchewan, Saskatoon, Saskatchewan, Canada S7N 0W0
† Université Catholique de Louvain, Groupe de Physico-Chimie, Minerale et de Catalyse, Louvain-la-Neuve 1348, Belgium

The biological availability and the toxicological effects of As depend on its chemical state[1,2]. It is therefore important to know the speciation and transformations of As in terrestrial and aquatic environments. Arsenite, As(III), is much more toxic than arsenate As(V) (ref. 2). The suspended and bottom sediments can potentially detoxify As(III) that enters aquatic systems by converting it to As(V) through abiotic oxidation[3]. We hypothesized that the oxides of Mn(IV) and Fe(III) may be the primary electron acceptors in the oxidation of As(III) because Mn and Fe readily participate in many oxidation–reduction reactions in natural environments; however, their role in the oxidation of As(III) to As(V) is yet to be established. We report here that Mn(IV) oxide is a very effective oxidant with respect to As(III); on the other hand, despite the thermodynamic favourability[3], the evidence obtained from colorimetry and X-ray photoelectron spectroscopy shows that a redox reaction between Fe(III) oxide and As(III) does not occur within 72 h, indicating that the kinetics of the redox reaction between As(III) and Fe(III) is relatively slow.

Birnessite (δ-MnO_2) was prepared and used in this study because it is one of the most common Mn oxides in sediments and soils[4]. The appearance of As(V) in solution after adding solutions of various concentrations of As(III) to Mn(IV) oxide (Table 1) clearly shows that As(III) is converted to As(V) by Mn(IV); in a control experiment, no detectable As(III) was oxidized in the absence of Mn(IV) oxide. With increasing additions of As(III) above 300 μg ml^{-1}, the amount of As(V) produced increased very little (Table 1), indicating that the oxidation capacity of the 100 mg sample of Mn(IV) oxide for As(III) was being approached. Assuming that the Mn(IV) oxide (birnessite) contains 9.5% K and the oxygen : manganese ratio is close to 2 (ref. 4), the 216 μg As(V) ml^{-1} in solution after adding 1,000 μg As(III) ml^{-1} to 100 mg of Mn(IV) oxide (Table 1) indicates that at least ~19% of the Mn(IV) initially present is reduced to Mn(II).

Mn(II) is more soluble than Mn(IV) (ref. 5), therefore the high concentrations of Mn and K in solution in the As(III)–Mn(IV) oxide systems relative to the As(V)–Mn(IV) oxide system (Table 1) is further evidence that Mn(IV) is reduced to Mn(II) by As(III). The reduction of Mn(IV) and the subsequent partial dissolution of the Mn oxide on the addition of As(III) disrupts the crystal structure at the surface of the oxide and thus a portion of the K present within the structure is free to pass into solution. As the Mn oxide contains ~9.5% K (ref. 4), the 90.9 μg K ml^{-1} in solution for the 1,000 μg As(III) ml^{-1} treatment (Table 1) represents ~67% of the K initially present in the Mn oxide sample. The 50 μg K ml^{-1} in solution in the As(V)–Mn oxide system (Table 1) is attributed to exchangeable K present at or near the surface of the oxide. The K in solution in the As(III)–Mn oxide systems is thus attributed to exchangeable plus structural K.

The decrease in the Mn concentration in solution with the increase in the initial As(III) concentration from 300 to 1,000 μg ml^{-1} (Table 1) is attributed to an increase in the formation of a sparingly soluble manganese–arsenate complex, $Mn_3(AsO_4)_2$. This interpretation is substantiated by the fact that the solubility product of $Mn_3(AsO_4)_2$ (ref. 6) is very much lower than the ion product, $(Mn^{2+})^3(AsO_4^{3-})^2$, of the equilibrated solutions in the As(III)–Mn(IV) oxide systems.

Manganese oxide sorbs little As(V) (Table 1). The point of zero charge of birnessite is close to *p*H 2 (ref. 7), and therefore the oxide has a high negative charge density at *p*H values used in this study; As(V) (largely $H_2AsO_4^-$ and $HAsO_4^{2-}$) is also negatively charged at the *p*H of this system[8], consequently a high energy barrier must be overcome for As(V) to sorb onto Mn oxide.

When freshwater lake sediments from southern Saskatchewan, Canada were treated with hydroxylamine hydrochloride or sodium acetate (effective extractants for the selective removal of Mn from sediments and soils[9,10]), the oxidation of As(III) to As(V) by the treated sediments was greatly decreased relative to untreated samples of the sediments[11]. This is supporting evidence that Mn oxide is a primary sediment component responsible for the oxidation of As(III). In sediments and soils Mn oxides occur as discrete particles and/or as coatings on other sediment/soil components[12]; and Mn is present in both the colloidal and non-colloidal particle size fractions of stream and lake sediments[10].

No detectable As(V) was found in solution after the addition of As(III) to Fe(III) oxide (Table 1), showing that Fe(III) oxide does not oxidize As(III) at neutral *p*H values within 72 h (earlier studies[3] showed that 48 h was sufficient time for 0.7 mg As(III) to disappear when added to 1 g of sediment). The data in Table 1 also show that As(III) is more strongly sorbed by Fe(III) oxide than is As(V); this is attributed to structural and charge differences between the two As complexes: As(III) (H_3AsO_3, coordination geometry is pyramidal, at *p*H values <9.2 arsenious acid is predominantly a neutral complex[8]) and As(V) (H_3AsO_4, coordination geometry is tetrahedral, at *p*H values >2.3 arsenic acid is largely dissociated and negatively charged[8]).

Table 1 Oxidation of As(III) and sorption of As by Mn(IV) and Fe(III) oxides

Mn(IV) oxide						Fe(III) oxide			
As(III) or As(V) added (μg ml^{-1})	As(III)	As(V) (μg ml^{-1} in solution)	Mn (μg ml^{-1} in solution)	K	Final *p*H	As(III) or As(V) added (μg ml^{-1})	As(III) (μg ml^{-1} in solution)	As(V) (μg ml^{-1} in solution)	Final *p*H
100 As(III)	ND	83.5±1.4*	0.41±0.12	60.5±0.0	7.1	100 As(III)	3.67±0.24	ND	6.9
300 As(III)	63.2±7.0	186±5	8.08±0.36	90.0±1.5	7.1	100 As(IV)	ND	53.6±1.1	7.7
500 As(III)	213±4	205±1	6.06±0.60	90.5±0.6	7.3				
1000 As(III)	665±5	216±4	4.61±0.86	90.9±0.8	7.5				
300 As(V)	ND	298±1	0.06±0.02	50.0±0.9	7.5				

The Mn(IV) (birnessite) and Fe(III) oxides were prepared as outlined elsewhere[4,15]. 100 mg of Mn(IV) or Fe(III) oxide was suspended in 70 ml of a solution containing the various concentrations of As(III) and As(V) as $NaAsO_2$ and $Na_2HAsO_4 \cdot 7H_2O$, respectively. The *p*H of the suspensions was adjusted to 7.5, and the flasks containing the suspensions were placed in an oscillating shaker at 25±0.2°C for 72 h. After the reaction period, the concentration of As(III) and As(V) in solution was determined colorimetrically as described elsewhere[3].

* Mean ±s.d.; *n* = 3. ND, not detectable.

0028-0836/81/190050—02$01.00

Table 2 Binding energies for As and Fe(III) reference oxides and As(ads)–Fe(III) oxide

Sample	$As_{3p_{3/2}}$	$As_{3p_{1/2}}$	Δ*	$Fe_{2p_{3/2}}$
$NaAsO_2$	143.4	148.2	4.8	–
$Na_2HAsO_4 \cdot 7H_2O$	144.1	148.9	4.8	–
Fe(III) oxide				711.7
As(III)†–Fe(III) oxide	143.7	148.4	4.7	711.3
As(V)†–Fe(III) oxide	144.4	149.2	4.8	711.5

The measurements were performed on a Vacuum Generators ESCA 3 Spectrometer equipped with a Tracor Northern NS 570 signal averager for signal-to-noise ratio enhancement. The samples were ground and dusted on double-sided adhesive tape. The sequence of spectra was such as to embrace the As_{3p} lines by the C_{1s} reference line at 285.0 eV. Each sample was run twice and the precision of the binding energies was ±0.1 eV. The signal-to-noise ratio was 50:1 for the C_{1s} and Fe_{2p} lines and at least 15:1 for the As_{3p} doublet.

* BE ($As_{3p_{1/2}} - As_{3p_{3/2}}$).

† Initial concentration of As(III) and As(V) was 100 μg ml^{-1} and the concentration of As(III) and As(V) remaining in solution after the 72-h reaction period is given in Table 1.

About 96% of the As(III) added to the Fe(III) oxide suspension is sorbed onto the surface of the oxide, and the colorimetric method yields no information on the oxidation state of the surface-sorbed As. X-ray photoelectron spectroscopy (XPS) is a powerful tool for investigating the oxidation state of surface-sorbed species. The results of an XPS study of the As–Fe(III) oxide system is presented in Table 2. The absolute use of binding energies alone for the assignment of the oxidation state is generally not reliable, especially in the case of an adsorbed species when the reference used is a bulk compound (in this case $NaAsO_2$ and $Na_2HAsO_4 \cdot 7H_2O$). However, the constancy of the 0.7-eV shift observed between the As(III) and As(V) samples ($As_{3p_{3/2}}$ and $As_{3p_{1/2}}$ lines) on one hand and between the As(III)–Fe(III) oxide and As(V)–Fe(III) oxide samples on the other (Table 2) is strong evidence that As(III) and As(V) are sorbed as such and no change in oxidation state occurs. The 0.3-eV increase when either As(III) or As(V) is absorbed onto iron oxide reflects extra atomic relaxation and/or Madelung potential effects. The inability of Fe(III) oxide to convert As(III) to As(V) within 72 h despite the thermodynamic favourability indicates that the kinetics of the redox reaction is relatively slow at pH values near neutrality.

The biological methylation of As, Hg and Se has attracted a great deal of attention in recent years. Because the methylation of inorganic arsenicals is essentially a reduction reaction[13], the abiotic oxidation of As(III) to As(V) by freshwater lake sediments would tend to counteract the methylation process. This is significant because some organic arsenicals such as the various methyl-substituted arsines are more toxic than the inorganic arsenic species in higher oxidation states[14].

Our data indicate that Fe(III) oxide does not convert As(III) to As(V) within 72 h. On the other hand, Mn(IV) oxide, which is common in freshwater sediments, is a very effective oxidant with respect to As(III). Our data thus indicate that Mn(IV) oxide may be very important in natural systems in decreasing the concentration of As(III), a highly toxic pollutant.

This study was supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

Received 6 February; accepted 6 March 1981.

1. Webb, J. L. *Enzyme and Metabolic Inhibitors* Vol. 3, Ch. 6 (Academic , New York, 1966).
2. Ferguson, J. F. & Gavis, J. *Wat. Res.* **6,** 1259–1274 (1972).
3. Oscarson, D. W., Huang, P. M. & Liaw, W. K. *J. Envir. Qual.* **9,** 700–703 (1980).
4. McKenzie, R. M. *Miner. Mag.* **38,** 493–502 (1971).
5. Stumm, W. & Morgan, J. J. *Aquatic Chemistry*, 547-553 (Wiley–Interscience, New York, 1970).
6. Hess, R. E. & Blanchar, R. W. *J. Soil Sci. Soc. Am.* **40,** 847–852 (1976).
7. Healy, T. W., Herring, A. P. & Fuerstenau, D. W. *J. Colloid Interface Sci.* **21,** 435–444 (1966).
8. Weast, R. C. (ed.) *CRC Handbook of Chemistry and Physics* 58th edn, D-151 (CRC, West Palm Beach, 1978).
9. Chao, T. T. *Proc. Soil Sci. Soc. Am.* **26,** 764–768 (1972).
10. Oscarson, D. W., Rogers, J. S., Huang, P. M. & Liaw, W. K. *Int. Rev. ges. Hydrobiol.* **66,** 95–107 (1981).
11. Oscarson, D. W., Huang, P. M. & Liaw, W. K. *Clays Clay Miner.* (in the press).
12. Jenne, E. A. *Am. Chem. Soc., Adv. Chem. Ser.* **73,** 337–387 (1968).
13. Braman, R. S. in *Arsenical Pesticides* (ed. Woolson, E. A.) 108–123 (American Chemical Society, Washington DC, 1975).
14. Penrose, W. R. *CRC Crit. Rev. Envir. Control.* **4,** 465–482 (1974).
15. Huang, P. M., Wang, T. S. C., Wang, M. K., Wu, M. H. & Hsu, N. W. *Soil Sci.* **123,** 213–219 (1977).

Earthquake-induced ground acceleration

Michele Caputo

Istituto di Fisica, G. Marconi, University of Rome,
Piazzale Aldo Moro, 2, I-00185 Rome, Italy

Earthquakes may be characterized after the event by several physical parameters, including the linear dimension of the fault, l, and the stress drop, p. Although the direct measurement of these parameters is often difficult, there are data on the frequency distribution of l and p for certain series of earthquakes (refs 1–11; C. Archambeau, unpublished data; personal communications from T. Mikumo, T. Utsu and R. I. Archuleta). For p there is evidence[12–14] that the frequency distribution $N(p)$ is given by a relationship of the form

$$\log N(p) = c_1 - (1-\alpha)\log p \qquad (1)$$

where c_1 and α are constants (α lying between 0 and 1) and $N(p)\,dp$ is the number of stress drop in the range p to $p + dp$ and $p_1 < p < p_2 (p_1 > 0$, where p_2 is limited). We report here that a relationship of this type holds (with specific parameters c_1 and α) for 15 different sets of earthquake occurring in Japan, California, Mexico, Nevada, Italy and the Aleutian arc. We also show that the frequency distribution $N(p)$ provides some information about the frequency distribution of ground acceleration during earthquakes.

Figure 1 verifies that the density distribution of the stress drops is a decreasing function of p, and also that equation (1) is a satisfactorily representation of the frequency distribution of p. The stress drops computed by Archambeau (unpublished data) for the Alaskan–Aleutian region and the Japanese region have been obtained by fitting observed seismograms with synthetic ones obtained from source models of the events that have the observed source parameters. The stress drops of the data of Mikumo and Utsu (personal communications) have been computed by various methods including Brune's formula[15]. The value of $-1+\alpha$ obtained from the analysis of Archambeau's data for Japan is -1.2, that from the data of Mikumo and Utsu is -1.4: considering that there are relatively few data in both sets these two values agree well.

Few authors report the accuracy of their data. However, Peppin and Bufe[9] computed the seismic moment and the stress drop for 40 earthquakes; they did not find large differences between the stress drops measured from vertical P-wave and horizontal S-wave spectral corner frequencies. They also computed the moment and stress drop using records from several different instruments. The scatter of the data is acceptable, 65% of their corner frequencies have precision better than 10%.

The regression lines of Fig. 1 have been computed using the method of the maximum likelihood assuming $\sigma(\log N)/\sigma(\log p) = 1$. The values of $1-\alpha$ are affected by a mean standard deviation of at most 10%.

To obtain the density distribution of the r.m.s. acceleration of the ground we also need the density distribution $N(l)$ of the linear dimension of the faults l. This has the form $N(l) = c_2 l^{-\nu}$ (ref. 12) with $\nu = 2b + 1$, where b is given by the results from the Gutenber-Richter frequency (n) magnitude (M) relation $\log n = a - bM$, b is typical of each seismic region and generally varies from 0.5 to 1.

0028–0836/81/190051—03$01.00

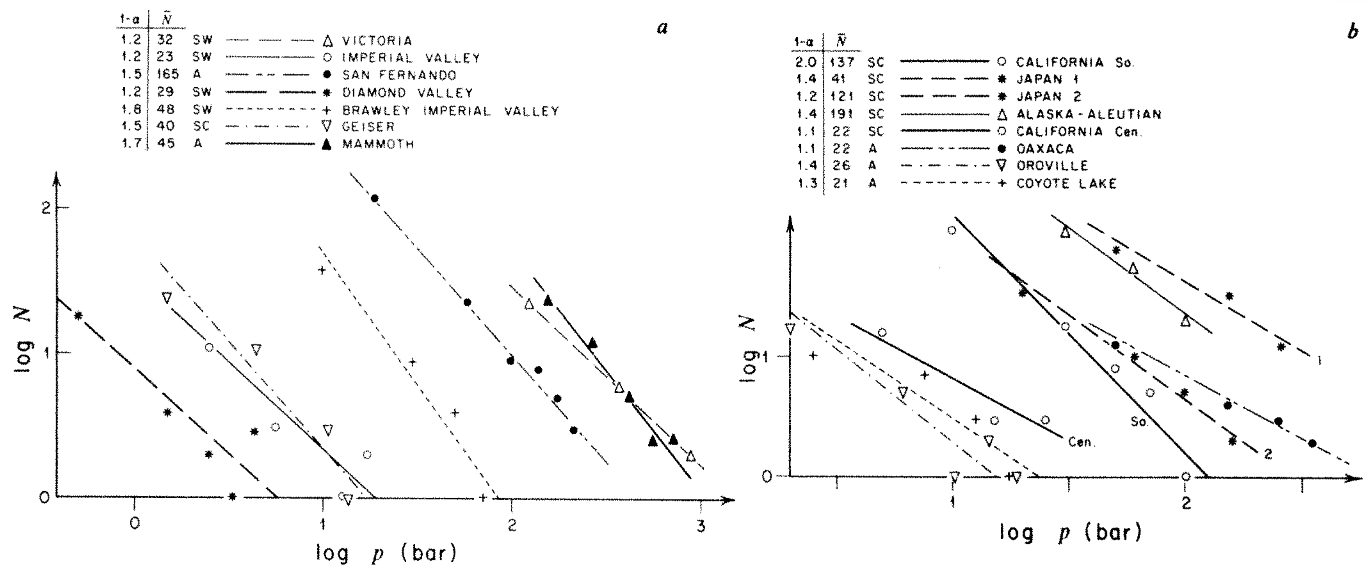


Fig. 1 Linear regression for the stress drops of aftershocks of: *a*, the 1971 San Fernando earthquake[11]; earthquakes of the Brawley–Imperial Valley 1975 swarm[5]; earthquakes of the Imperial Valley 1977 swarm[1]; scattered earthquakes in the Geiser (California) region[9]; earthquakes of the 1978 Victoria (Mexico) swarm[8]; earthquakes of a sequence in Diamond Valley (Sierra Nevada)[3]; aftershocks of the May 1980 Mammoth earthquake (R. I. Archuleta, personal communication), $\bar{N}$ is the total number of events used in the analysis. SC, Scattered events; A, aftershocks; SW, swarm. *b*, The 1979 Coyote Lake (California) earthquake[6]; aftershocks of the 1978 Oroville (California) earthquake[4]; earthquakes of the 1974 sequence of Central California[2]; aftershocks of the 1978 Oaxaca (Mexico) earthquake[7]; scattered earthquakes in the Aleutian and Alaskan regions (C. Archambeau, unpublished data); scattered earthquakes in Japan (1) (C. Archambeau, unpublished data); scattered earthquakes in Japan (2) (personal communications from T. Mikumo, and T. Utsu); scattered earthquakes in Southern California[10].

If the density distributions of p and l are known the density distribution of the r.m.s. ground acceleration a, can be calculated using the relationship established by Mcguire and Hanks[17]

$$a = 0.317\frac{p}{\rho R}\left(\frac{f_{\max}}{f_0}\right)^{1/2} \tag{2}$$

where ρ is the density, R the hypocentral distance, f_0 the spectral corner frequency of the seismic signal and $f_{\max} = Q\beta/\pi R$ is determined by the distance from the point to the epicentre R, the velocity of the waves β and the specific dissipation Q. Substituting[15]

$$f_0 = 2.34\beta/2\pi l$$

in equation (2) we find

$$a = \frac{0.29}{\rho}(Q)^{1/2}\frac{p}{R}\left(\frac{l}{R}\right)^{1/2} = kpl^{1/2}, \quad k = \frac{0.52}{\rho R}\left(\frac{f_{\max}}{\beta}\right)^{1/2} \tag{3}$$

The cumulative distribution function $n(a)$ of a can be determined from the density distributions (1) of l and p as follows. If p_1, p_2, l_1, l_2 are the minimum and maximum values of p and l, of the seismic region[16] then by integrating[12] in the domain of the plane l, p defined by the minimum and maximum values of p and l, we find that the cumulative distribution of a for $a > kp_1\sqrt{l_2}$ (or $a > kp_2\sqrt{l_1}$ whichever is larger) is

$$n(a) = \frac{D}{1-\nu}\left\{\frac{p_2^{\alpha}l_2^{1-\nu}}{\alpha} - \frac{(2\nu-2)l_2^{1-\nu-\alpha/2}a^{\alpha}}{\alpha(2\nu+\alpha-2)k^{\alpha}} - \frac{p_2^{2\nu+\alpha-2}a^{2-2\nu}}{k^{2-2\nu}(2\nu+\alpha-2)}\right\} \tag{4}$$

In the crust it is reasonable to assume $\rho = 2.8$ and $Q = 300$ for the seismic waves of interest, this gives $k = 1.81R^{-3/2}$. Because $\alpha < 0$, $\nu > 2$ it is seen that $n(a)$ is an increasing function of p_2 and decreasing function of a and R.

For small values of a in the range $p_1k\sqrt{l_2}$ to $p_2k\sqrt{l_1}$ such as those considered here the second term of equation (4) is small relative to the third. Because $\nu \simeq 2.8$, $n(a)$ is nearly proportional to $(aR^{3/2})^{-\delta}$ $(\delta \simeq 3.5)$.

Note that, considering equation (3) and the density distributions of p and l given by equation (1), a given pair of values of a and R, correspond to an infinite number of values of M and that, of all the values of M associated with this pair, Fig. 2, the minimum values of M is that which has the highest probability of occurrence. This implies that the return periods of the accelerations of the ground caused by earthquakes are much longer than previously estimated through empirical relations between M and a.

If we take into account that the seismicity is not concentrated in a point at distance R but covers an area, which we assume here for simplicity to be a sector of a circular ring defined by the angles ϑ_1, ϑ_2 and by the radii R_1 and R_2, then $n(a)$ is

$$n(a) = \frac{Dp_2^{\alpha}l_2^{1-\nu}}{1-\nu}\left\{\frac{1}{\alpha} - \frac{4(\nu-1)(a/kp_2\sqrt{l_2})^{\alpha}(R_3^{\frac{3}{2}\alpha+2} - R_1^{\frac{3}{2}\alpha+2})}{(\frac{3}{2}\alpha+2)(2\nu+\alpha-2)(R_2^2-R_1^2)\alpha} - \frac{2(a/kp_2\sqrt{l_2})^{2-2\nu}(R_3^{5-3\nu}-R_1^{5-3\nu})}{(2\nu+\alpha-2)(5-3\nu)(R_2^2-R_1^2)}\right\}$$

where $R_3 = R_2$ or $(p_2\sqrt{l_2})^{2/3}$ whichever is smaller.

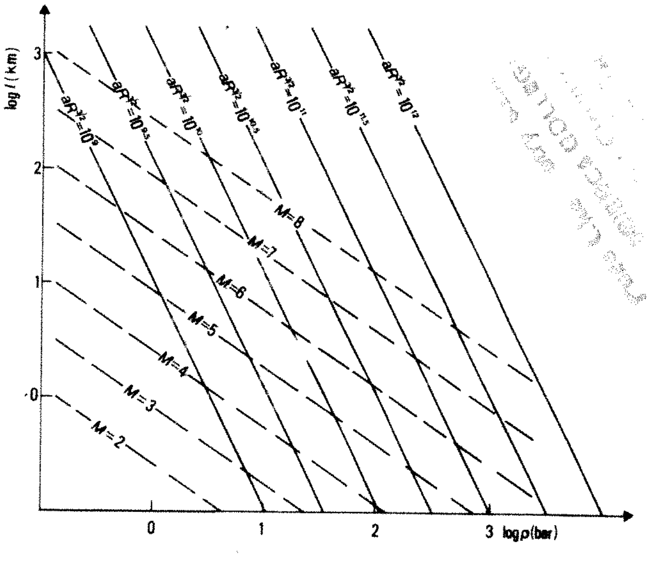


Fig. 2 Isolines of magnitudes and r.m.s. acceleration a in the plane of the linear dimensions of faults l and stress drops p. The lines $aR^{\frac{1}{2}}$ = constant have been computed from equation (3), the lines M = constant have been computed according to the formula $10^{11.1+1.46M} = \frac{16}{7}(l^3p^2/2\mu)$ of ref. 13.

In California $\nu = 2.78$, $\alpha = -1$, (ref. 13); a reasonable value of p_2 is 500 bar (refs 10, 13); we may tentatively consider a maximum linear dimension of a fault $l_2 = 5 \times 10^6$. Using equation (15) of ref. 12 with $p_2 = 5 \times 10^8$, $p_1 = 5 \times 10^5$ and the data of Epstein and Lomnitz[18] on Northern California Earthquakes one obtains $D = 10^{16.35}$ (ref. 12). For $a = 50$ Gal, $R_2 = 8R_1 = 8 \times 10^6$, equation (4) gives $n(50) = (17)^{-1}$ or a return period of 17 yr; this is much longer than the return period of $M = 5$ which, as shown by Fig. 2, could cause an acceleration of 50 Gal in the rare case when $l = 100$ m, and $p = 500$ bar.

Applying equation (2) to Southern California[17] and Friuli (E. Faccioli, personal communication) gave r.m.s. accelerations of the ground close to those observed; however, a test on a sequence of earthquakes recently observed in Northern California gave ambiguous results (T. C. Hanks, personal communication). If for some regions a formula for a must be used that is different from equation (2), such as that of Bernreuter[19] $a = (0.19p/\rho R)(l/R)^{0.7}$, the present method would still be valid, and a slightly different form for $n(a)$ would be obtained.

Received 19 November 1980; accepted 25 February 1981.

1. Adair, R., Brune, J. N. & Orcutt, J. *Earthquake Not.* **49,** 4 (1978).
2. Bakun, W. H., Bufe, G. C. & Stewart, R. M. *Bull. seism. Soc. Am.* **66,** 363 (1976).
3. Sommerville, M. R., Peppin, W. A. & Van Wormer, J. R. *Bull. seism. Soc. Am.* **70,** 1547 (1980).
4. Fletcher, J. B. *Bull. seism. Soc. Am.* **70,** 735 (1980).
5. Hartzell, S. H. & Brune, J. N. *Pageophysics* **115,** 333 (1977).
6. Lee, W. H. K., Herd, D. G., Cagnetti, V., Bakun, W. H. & Rapport, A. *U.S. geol. Surv. Open File Rep.* **79,** 1621 (1979).
7. Munguia, L. *et al. Geofis. Int.* **17,** 359 (1978).
8. Munguia, L. *et al. EOS* **59,** 1131 (1978).
9. Peppin, W. A. & Bufe, C. *Bull. seism. Soc. Am.* **70,** 269 (1980).
10. Tatcher, A. & Hanks, T. C. *J. geophys. Res.* **78,** 8547 (1973).
11. Tucker, B. E. & Brune, J. N. *Geol. Geophys. Stud.* **3,** 65 (1973).
12. Caputo, M. *Ann. Geofis.* **4,** 277 (1976).
13. Caputo, M. & Console, R. *Tectonophysics* **67,** T13 (1980).
14. Caputo, M. *Tectonophysics* **69,** T9 (1980).
15. Brune, J. N. *J. geophys. Res.* **75,** 4997 (1970).
16. Hanks, T. C. *J. geophys. Res.* **84,** 2235 (1979).
17. Mcguire, R. K. & Hanks, T. C. *Bull. seism. Soc. Am.* **5,** 70 1910 (1980).
18. Epstein, B. & Lomnitz, C. *Nature* **211,** 1197 (1966).
19. Bernreuter, D. L. *OECD-NEA/CSNI Rep.* No. 28, **2,** 429 (1979).

Can Kimberlites be generated from an ordinary mantle?

J. D. Kramers*, C. B. Smith†, N. P. Lock‡, R. S. Harmon§ & F. R. Boyd‖

* Department of Earth Sciences, University of Leeds,
Leeds LS2 9JT, UK
† Bernard Price Institute of Geophysical Research, University of The Witwatersrand, Johannesburg, South Africa
‡ Department of Geology, University of Sheffield,
Sheffield S10 2TN, UK
§ Scottish Universities Research and Reactor Centre, East Kilbride,
Glasgow G75 0QU, UK
‖ Geophysical Laboratory, 2801 Upton Street, NW,
Washington, DC 20008, USA

Kimberlites are characterized by a high concentration of incompatible elements, including light rare earths. However, Nd isotopic evidence indicates that their source regions do not have a long history of enrichment. If kimberlites can be generated from non-enriched mantle sources by simple partial melting processes, this implies that crystal–liquid partition coefficients for some trace elements are lower for kimberlitic liquids than for basalts or andesites. From a comparison of clinopyroxene megacrysts (regarded as equilibrated with kimberlite at depth) with existing kimberlite data we argue here that low crystal–liquid partition coefficients for the rare earths are plausible.

Experimental work on the melting relations of model peridotitic systems charged with CO_2 has shown that liquids of carbonatitic and kimberlitic composition may be produced by a small degree of partial melting of a carbonated peridotite[1,2]. This melting is likely to be caused by a change in the P–T conditions rather than an incursion of CO_2, because CO_2 cannot exist as a free fluid phase in a silicate mantle at pressures higher than ~32 kbar, but is bound by the formation of carbonate minerals[2]. The melting constitutes a single process and, therefore, trace-element enrichment in the liquid relative to its source is limited by solid–liquid partition coefficients.

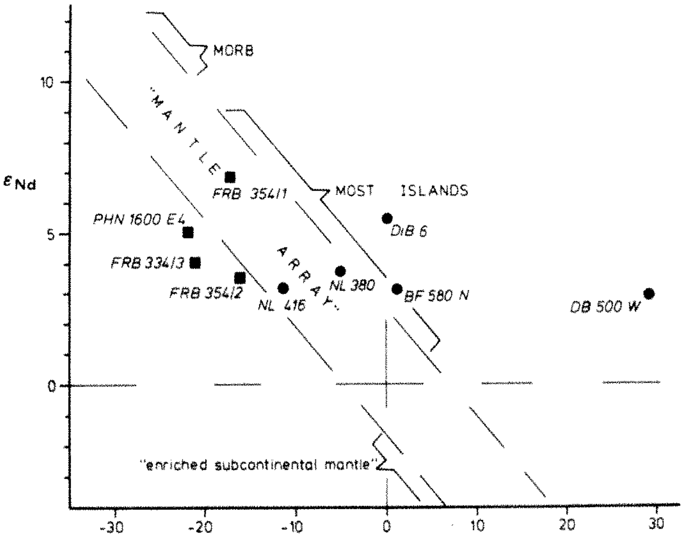


Fig. 1 ε_{Nd} plotted against ε_{Sr} showing kimberlite (●) and diopside megacryst (■) samples. The age[37] used for correction is 90 Myr. The 'mantle array' contains data from mid-ocean ridge basalts (MORB) and most oceanic islands[5] as well as 'enriched subcontinental mantle' as exemplified by diopsides from peridotite nodules in kimberlite[9].

$$\varepsilon_{Nd} = \left[\frac{(^{143}Nd/^{144}Nd)_{init} \text{ of sample}}{(^{143}Nd/^{144}Nd)_t \text{ of CHUR}} - 1 \right] \times 10^4$$

where t denotes the age of the sample. Analogous for ε_{Sr}. $(^{147}Sm/^{144}Nd)_{CHUR} = 0.194$; $(^{143}Nd/^{144}Nd)_{CHUR}$(present day) = 0.51264; $(^{87}Rb/^{86}Sr)_{UR} = 0.0827$; $(^{87}Sr/^{86}Sr)_{UR}$(present day) = 0.7045.

An assessment of the time-integrated enrichment or depletion in incompatible elements, including light rare earths (LREs), in the sources of volcanic rocks can be made from Nd and Sr isotope data. $^{87}Sr/^{86}Sr$ ratios of fresh kimberlites are ~0.704 (ref. 3), indicating time integrated source Rb/Sr ratios lower than that of the 'uniform reservoir' (UR)[4] also known as 'bulk earth'[5]. Recent Sm–Nd work on kimberlites ranging in age between 1,300 and 90 Myr (ref. 6) seems to point to time-integrated source Sm/Nd ratios close to that of the 'chondritic uniform reservoir' (CHUR)[4] or 'bulk earth'[5]. In contrast, combined Sr and Nd work on American kimberlites[7] has suggested both time integrated large-ion lithophile (LIL) and LRE element enrichment and depletion for their sources. We have measured the $^{143}Nd/^{144}Nd$ and $^{87}Sr/^{86}Sr$ ratios of five Cretaceous Southern African kimberlites, using methods described elsewhere[8], and the results are listed in Table 1. In the $\varepsilon_{Nd}/\varepsilon_{Sr}$ plot[4] (Fig. 1) the two dyke kimberlites from Letseng la Terai (NL 380 and 416) and the Bultfontein sample (BF 580 N) plot close to the 'mantle array' which is defined by both oceanic and continental volcanics[5] as well as some mantle derived xenoliths[9]. They indicate a slight time-integrated LIL and LRE element depletion of their sources. Kimberlite samples DiB 6 and DB 500 W plot well to the right of the mantle array, which is thought to be the result of Sr exchange with circulating groundwater[3]. Any effect on the $^{143}Nd/^{144}Nd$ ratios is likely to have been a lowering one[4]; thus these two kimberlites were probably also derived from sources with a time integrated LREE depletion. From published Sr and Nd isotopic data on kimberlites[3,6,7] and our results, the source regions of kimberlites seem to be similar in their LIL and LRE element characteristics to the sources of ocean island volcanics. These are characterized

0028–0836/81/190053—04$01.00

Table 1 Isotopic data on kimberlites and clinopyroxene megacrysts

Sample	$^{87}Sr/^{86}Sr$	$^{143}Nd/^{144}Nd$	$^{87}Rb/^{86}Sr$	$^{147}Sm/^{144}Nd$	$\delta^{18}O$ (SMOW)
Kimberlites					
NL 380 (Dyke, Letseng la Terai, Lesotho)	0.70439±3	0.51278±4	0.281	0.0936	
NL 416 (Dyke, Letseng la Terai, Lesotho)	0.70386±2	0.51274±3	0.218	0.0916	
DiB 6 (Wesselton Mine, Kimberley, South Africa)	0.70456±6	0.51285±2	0.127	0.0717	
BF 580 N (Bultfontein Pipe, Kimberley, South Africa)	0.70461±3	0.51274±2	0.119	0.09	
DB 500 W (De Beers Pipe, Kimberley, South Africa)	0.70691±10	0.51273±2	0.353	0.915	
Megacrysts					
FRB 334/3 (Camutue Pipe, Angola)	0.70293±20	0.51283±3	0.0082	0.147	+6.57
PHN 1600 E4 (Thaba Putsoa Pipe, Lesotho)	0.70280±20	0.51289±3	0.0023	0.168	+6.56
FRB 354/1 (Kimberley Pool, Kimberley, South Africa)	0.70317±18	0.51302±4	0.0026	0.191	+6.35
FRB 354/2 (Kimberley Pool, Kimberley, South Africa)	0.70326±14	0.51281±2	0.004	0.168	
FRB 354/3 (Kimberley Pool, Kimberley, South Africa)					+6.88
FRB 354/4 (Kimberley Pool, Kimberley, South Africa)					+6.50
IM 7-C1 (Front Range kimberlites, Wyoming)	0.70360±10		0.003		
Standards					
NBS 987 Sr (66 runs since 1977)	0.710272±38 (1 s.d. of population)				
La Jolla Nd standard (45 runs since 1977)		0.511880±36 (1 s.d. of population)			

Errors quoted on Sr and Nd isotope ratios are 2 standard deviations of the mean (except for standards). Errors on Rb/Sr and Sm/Nd ratios estimated at 2%. $^{143}Nd/^{144}Nd$ ratios normalized to $^{146}Nd/^{144}Nd = 0.7219$.

by trace element and isotopic heterogeneity, varying from no differentiation (relative to CHUR) to moderate LIL and LRE element depletion[5]. They also exhibit the same heterogeneous, mostly radiogenic lead isotopic signature as kimberlites[10].

Geochemical work on kimberlites[11-13] has highlighted their strong enrichment in incompatible elements, including LREE. One explanation invoked for this is that their mantle sources were pre-enriched[13,14], but long-term enrichment of kimberlite sources can be discounted on the grounds of the isotopic evidence discussed above, with the possible exception of highly micaceous kimberlites, which can have genuinely high initial $^{87}Sr/^{86}Sr$ ratios[3]. From isotopic[9,10,15] and trace element[16-18] studies on mantle derived xenoliths it has been shown that provinces exist in the upper mantle underneath Southern Africa which are anomalously enriched in LIL and LRE elements and have been so for a long period of geological time[9]. However, these were found to occur at shallower levels than the source regions of kimberlites[19,20] and their high $^{87}Sr/^{86}Sr$ ratios preclude a genetic link with most kimberlites.

The idea that kimberlites, with their strong enrichment in incompatible elements, could be derived from an 'ordinary mantle' which was slightly depleted in these elements, is an apparent paradox. Mechanisms which could resolve this paradox include:

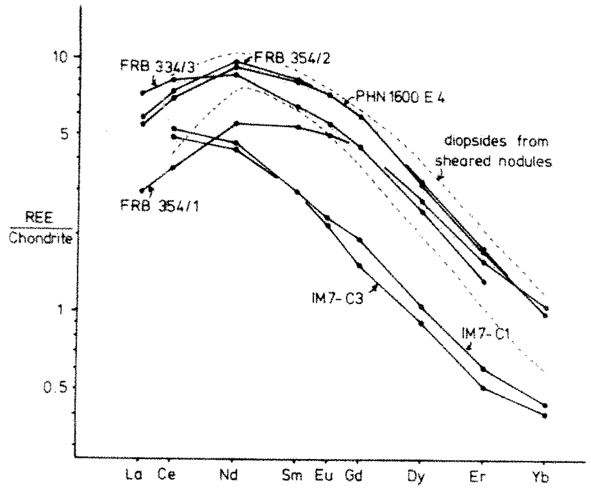


Fig. 2 Chondrite normalized rare earth abundances in diopside megacrysts. Stippled area, field for diopsides from Northern Lesothan sheared lherzolite nodules analysed by Shimizu[17].

(1) Transport of some incompatible elements as complexes in a CO_2 or H_2O vapour phase[11]. This mechanism may account for the great heterogeneities found in diatremes at near-surface level, but is unlikely to operate at the depth of origin of kimberlites[1,2].

(2) Metasomatic enrichment of the kimberlite source region shortly before melting, as proposed for some alkali basalts[21]. There is no evidence excluding this process, but for kimberlites the metasomatizing agent would have to be a melt[1,2] and the mechanism therefore involves 'two-stage' melting.

(3) Simple melting with trace-element partitioning governed by lower solid–liquid partition coefficients than are found in basaltic systems. This alternative is examined below, using K, Rb, Sr and REE data on subcalcic clinopyroxene megacrysts.

Among the otherwise variable inclusion population in kimberlite pipes, megacrysts (monomineralic inclusions over ~5 mm in size) of subcalcic clinopyroxene, garnet, olivine, orthopyroxene, phlogopite and ilmenite stand out by being almost ubiquitous. Subcalcic clinopyroxene megacrysts typically indicate equilibration temperatures[19,22,23] around 1,300–1,400 °C. Chemically they do not closely resemble clinopyroxenes in xenoliths (although some similarities exist with Northern Lesothan 'sheared lherzolites'[17,19]), and unlike other types of inclusions in kimberlites they consistently have low $^{87}Sr/^{86}Sr$ ratios[15]. They have been regarded as phenocrysts crystallized from the kimberlite magma at depth[22,24], which is particularly plausible for the low-Cr (0.3–0.6% Cr_2O_3) variety[22]. Because their composition usually indicates a whole range of temperatures of equilibration[19,22], this interpretation implies a diapiric origin for kimberlite[25], with fractional crystallization of phenocrysts occurring over a considerable depth range[22].

To test the isotopic consistency of the phenocryst hypothesis, we have measured $^{143}Nd/^{144}Nd$, $^{87}Sr/^{86}Sr$ and $^{18}O/^{16}O$ ratios of four subcalcic low-Cr diopside megacrysts from Southern Africa and the $^{87}Sr/^{86}Sr$ ratio of a similar diopside megacryst from the Colorado–Wyoming Front Range kimberlite province. Descriptions and major element analyses of the samples have been published elsewhere[19,22,23,26]. The isotopic results are listed in Table 1. The ε_{Nd} and ε_{Sr} values (Fig. 1) plot close to the mantle array in a 'moderately depleted' region. We have argued above that the kimberlite points in Fig. 1 were probably moved to the right (and possibly slightly downward) by interaction with groundwater, and might originally have plotted close to the mantle array in a similarly depleted region. The $\delta^{18}O$ values of the megacrysts (Table 1) are uniform around +6.5, a value higher than that obtained for most ultramafic xenoliths, but similar to phlogopite megacrysts in kimberlite[27]. Lead isotope ratios of the four Southern African clinopyroxene megacrysts (J.D.K., unpublished data) plot within the field defined by Southern African Cretaceous kimberlites[10].

Although the isotopic data discussed above do not prove that subcalcic clinopyroxene megacrysts crystallized from the kimberlite magma at depth, they do not conflict with this hypothesis, which has also stood up to petrological scrutiny[22]. Below, assuming equilibrium as a working hypothesis, we derive clinopyroxene–kimberlitic liquid partition coefficients for K, Rb, Sr and the REE and compare them with literature values for basaltic systems. Rb, K, Sr and REE concentrations measured by isotope dilution are listed in Table 2, including results on the kimberlite and megacryst samples listed in Table 1 and one additional diopside megacryst from the Front Range kimberlites. The chondrite normalized REE patterns of the clinopyroxenes (Fig. 2) are similar to patterns found by Shimizu[17] for clinopyroxenes from Northern Lesothan 'sheared' lherzolite nodules, for which an approximately chondritic whole rock REE pattern was reconstructed from the mineral data[17]. The subcalcic clinopyroxene megacrysts could thus be in equilibrium with a non-enriched mantle, either directly or through a liquid.

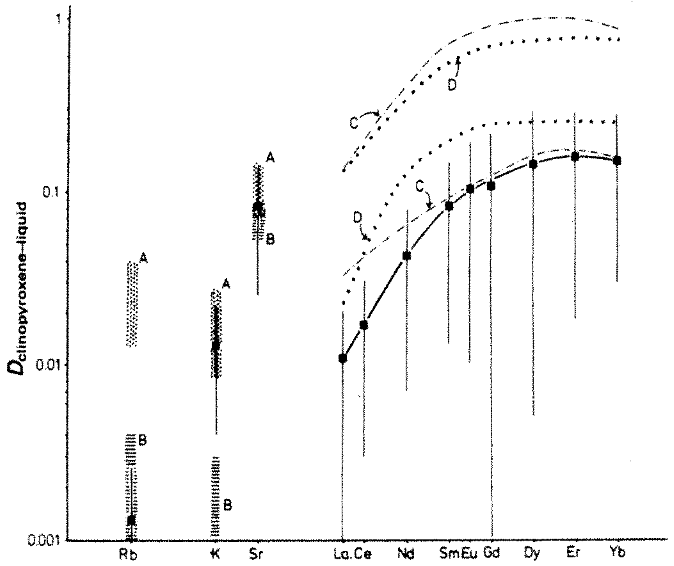


Fig. 3 Apparent clinopyroxene–kimberlitic liquid partition coefficients found in the present investigation (■, with one standard deviation error bars) compared with ranges of literature values for basaltic liquids. A, Philpotts and Schnetzler[33]; B, Shimizu[29]; C, boundaries of field for data of Onuma *et al.*[32] and Schnetzler and Philpotts[34], D, boundaries of field for data of Nicholls and Harris[31] and work reviewed by Irving[30].

The great heterogeneity of kimberlite occurrences and the variation between them means that the approach to the derivation of partition coefficients can only be semiquantitative, and we use an 'average kimberlite'. Its REE pattern is derived from the mean values of 23 kimberlite analyses, including the five listed in Table 2 and results from the literature[11–13,28], and its K, Rb and Sr concentrations are taken from values in Table 2 together with 11 analyses from the literature[10,15], avoiding upper diatreme facies kimberlites which can be heavily depleted in alkalies. The values for this 'average kimberlite' are also listed in Table 2, along with standard deviations. The apparent clinopyroxene–kimberlitic liquid partition coefficients listed in Table 2 have been derived from the average of the clinopyroxene analyses and the 'average kimberlite'. In Fig. 3 these apparent partition coefficients are compared with results obtained from experimental[29–31] and natural[32–34] systems. The value for Sr agrees well with the data from clinopyroxene–basaltic liquid studies[29,33]. For K and Rb there is a considerable discrepancy between Shimizu's[29] experimental results and Philpott's and Schnetzler's[33] data (Fig. 3). Our apparent partition coefficient for K agrees with the latter, whereas our value for Rb conforms to Shimizu's results. The apparent REE clinopyroxene–kimberlitic liquid partition coefficients follow a pattern which is strikingly similar in shape to those derived from both natural

Table 2 Trace element concentrations (p.p.m.) and apparent partition coefficients clinopyroxene–kimberlitic liquid

Sample	Rb	K	Sr	La	Ce	Nd	Sm	Eu	Gd	Dy	Er	Yb
Kimberlites												
NL 379 (Letseng la Terai)	64	4441	932	79.6	154	62.0	9.83	2.65	6.61	3.50	1.21	0.74
NL 380	95	7972	977	168	324	125	19.3	4.95	12.4	6.77	2.53	1.91
NL 416	63	5894	836		193	74.7	11.3	2.88	7.04	3.51	1.23	0.75
DiB 6	120	16700	2727	258	465	190	22.6	6.92	20.5	9.30	2.90	1.62
DB 500 W	72		606	65.1	129	57.4	8.71	1.86	5.68	3.20	1.30	1.00
Megacrysts												
FRB 334/3	0.31	237	112	2.35	7.02	5.35	1.30	0.42	1.22	0.85	0.31	
PHN 1600 E4	0.07	184	89	1.77	6.00	5.80	1.61	0.54	1.60	1.08	0.40	0.22
FRB 354/1	0.06	173	70	0.98	3.16	3.45	1.09	0.38	1.22	0.93	0.37	0.23
FRB 354/2	0.12	228	89	1.89	6.24	6.00	1.67	0.55	1.64	1.10	0.41	0.23
IM 7-C1	0.15	249	126		4.45	2.90	0.60	0.18	0.53	0.36	0.14	0.10
IM 7-C3 (Wyoming)		332			4.21	2.75	0.60	0.17	0.41	0.31	0.12	0.09
average	0.14 ±0.10	233 ±57	97 ±22	1.75 ±0.57	5.18 ±1.46	4.37 ±1.50	1.15 ±0.47	0.37 ±0.17	1.10 ±0.52	0.77 ±0.35	0.29 ∓0.13	0.18 ±0.07
Average kimberlite (see text)	108 ±35	17600 ±8000	1175 ±550	165 ±92	310 ±160	102 ±50	14 ±6	3.6 ±1.6	10.4 ±6	5.3 ±2.7	1.83 ±0.8	1.20 ±0.5
Apparent partition coefficient clinopyroxene–kimberlite	0.0013 ±0.0013	0.013 ±0.009	0.083 ±0.058	0.011 ±0.010	0.017 ±0.014	0.043 ±0.036	0.082 ±0.069	0.103 ±0.093	0.106 ±0.111	0.145 ±0.140	0.158 ±0.140	0.150 ±0.121

Errors listed with averages are standard deviations of populations. Errors quoted for partition coefficients were calculated from these standard deviations.

basaltic systems[32,34] and experimental work[30,31] (Fig. 3), but the values are up to an order of magnitude lower. A general observation in studies on both synthetic and natural systems[30,31,34] has been that crystal–liquid partition coefficients decrease with decreasing SiO_2 and increasing MgO content (that is, decreasing polymerization) of the liquid, and kimberlites are much poorer in SiO_2 (30–40%) and richer in MgO (generally 20–30%) than any siliceous volcanic rocks used in partition coefficient studies. A decrease in polymerization of the melt and thus in partition coefficient values could also result from dissolved H_2O (ref. 31). Although dissolved CO_2 may not have a depolymerizing effect, it might lower REE partition coefficients by complexing with REE, as is in evidence in vapour-phase CO_2 (ref. 35). Finally, crystal–liquid partition coefficients generally decrease with increasing temperature[30,31]. Without attempting a quantitative assessment of these effects, it can be predicted from the physicochemical differences between a kimberlitic and a basaltic liquid in their respective source regions that REE crystal–liquid partition coefficients would be considerably lower for the former. The low partition coefficients found for the REE in this work would thus not be unreasonable. Partition coefficients for Sr seem to be less affected by the basicity of the melt than those for the REE[33], and the present results are consistent with this observation.

Because the low apparent values found for clinopyroxene–kimberlitic liquid REE partition coefficients, compared with basaltic systems, would be largely understandable as resulting from properties of the liquid rather than the solid, the REE partition coefficients for other minerals in equilibrium with a kimberlitic liquid could be expected to be correspondingly lowered. A kimberlite with a typical strongly LREE-enriched pattern could thus be in REE equilibrium with ordinary mantle material such as forms the source of ocean island basalts. However, equilibrium with the original parent material is only a boundary condition of any melting model and implies either an almost infinitely small percentage of partial melt, or complete equilibration of a relatively small amount of melt with a very large volume of rock similar to the source rock. This ultimately happens in zone refining[36], but could also be achieved if the melt existed as intergranular films in a source rock undergoing penetrative deformation, for instance, if melting occurred in a rising diapir[25].

Although the above argument is in some ways speculative, the hypothesis of kimberlite generation from an ordinary mantle by a single melting process merits consideration from a trace element point of view.

We thank J. C. Roddick for analytical assistance and W. Wilkinson for mineral purification. Isotope research at Leeds is supported by the NERC which is thanked by J.D.K. for a grant. P. J. Betton, R. A. Cliff, M. J. Norry and M. F. Thirlwall provided constructive criticism.

Received 17 November 1980; accepted 17 February 1981.

1. Eggler, D. H. *Yb Carnegie Instn Wash.* **74,** 468–474 (1975).
2. Wyllie, P. J. *J. Geol.* **86,** 687–713 (1978).
3. Barrett, D. R. & Berg, G. W. *Phys. chem. Earth* **9,** 619–636 (1975).
4. DePaolo, D. J. & Wasserburg, G. J. *Geochim. cosmochim. Acta* **43,** 615–628 (1979).
5. O'Nions, R. K., Carter, S. R., Evensen, N. M. & Hamilton, P. J. *A. Rev. Earth & planet. Sci.* **7,** 11–38 (1979).
6. Basu, A. H. & Tatsumoto. M. *Science* **205,** 398–401 (1979).
7. DePaolo, D. J. *U.S. Geol. Surv. Open File Rep.* 1978–701, 81–93 (1978).
8. Hawkesworth, C. J. *et al. Earth planet Sci. Lett.* **42,** 45–57 (1979).
9. Menzies, M. & Murthy, V. R. *Nature* **283,** 634–636 (1980).
10. Kramers, J. D. *Earth planet. Sci. Lett.* **34,** 419–431 (1977).
11. Fesq, H. W., Kable, E. J. D. & Gurney, J. J. *Phys. chem. Earth* **9**, 687–707 (1975).
12. Mitchell, R. H. & Brunfelt. A. O. *Phys. chem. Earth* **9,** 671–686 (1975).
13. Frey, F A., Ferguson, J. & Chappell, B. W. *2nd Int. Kimberlite Conf. Ext. Abstr.* (Carnegie Institution, Washington DC, 1977).
14. Kramers, J. D. *Earth planet. Sci. Lett.* **42,** 58–70 (1979).
15. Barrett, D. R. *Phys. chem. Earth* **9,** 637–654 (1975).
16. Shimizu, N. *Phys. chem. Earth* **9,** 655–669 (1975).
17. Shimizu. N. *Earth planet. Sci. Lett.* **25,** 26–32 (1975).
18. Nixon, P. H., Rogers, N. W., Gibson, I. L. & Grey, A. *A. Rev. Earth planet. Sci.* (in the press).
19. Boyd, F. R. & Nixon, P. H. in *Lesotho Kimberlites* (ed. Nixon, P. H.) 67–75 (Lesotho National Development Corporation, 1973).
20. Dawson, J. B., Gurney, J. J. & Lawless, P. J. *Nature* **257,** 299–300 (1975).
21. Menzies, M. & Murthy, V. R. *Earth planet. Sci. Lett.* **46,** 323–334 (1980).
22. Eggler, D. H., McCallum, M. E. & Smith, C. B. in *The Mantle Sample*, 213–226 (American Geophysical Union, 1979).
23. Boyd. F. R. & Nixon, P. H. *Geochim. cosmochim. Acta* **42,** 1367–1382 (1978).
24. Nixon, P. H., von Knorring, O. & Rooke, J. M. *Am. Miner.* **48,** 1090–1132 (1963).
25. Green, H. W. II & Gueguen, Y. *Nature* **249,** 617–620 (1974).
26. Boyd, F. R. & Danchin, R. V. *Am. J. Sci.* **280**-A (Jackson Vol.) 528–549 (1980).
27. Sheppard, S. M. F. & Dawson, J. B. *Phys. chem. Earth* **9,** 747–763 (1975).
28. Smith, C. B. thesis, Colorado State Univ., Fort Collins (1977).
29. Shimizu, N. *Geochim. cosmochim. Acta* **38,** 1789–1798 (1974).
30. Irving, A. J. *Geochim. cosmochim. Acta* **42,** 743–770 (1978).
31. Nicholls, A. & Harris, K. L. *Geochim. cosmochim. Acta* **44,** 287–308 (1980).
32. Onuma, N., Higuchi, H., Wakita, H & Nagasawa, H. *Earth planet Sci. Lett.* **5,** 47–51 (1968).
33. Philpotts, J. A. & Schnetzler, C. *Geochim. cosmochim. Acta* **34,** 307–322 (1970).
34. Schnetzler, C. & Philpotts, J. A. *Geochim. cosmochim. Acta* **34,** 331–340 (1970).
35. Wendlandt, R. F. & Harrison, W. J. *Contr. Miner. Petrol.* **69,** 409–419 (1979).
36. Harris, P. G. & Middlemost, E. A. K. *Lithos* **3,** 77–88 (1969).
37. Davis, G. L. *2nd Int. Kimberlite Conf. Ext. Abstr.* (Carnegie Institution, Washington DC, 1977).

Late Phanerozoic extent of dry land

J. Graham Cogley

Department of Geography, Trent University, Peterborough, Ontario, Canada K9J 7B8

The extent of dry land during post-Carboniferous time seems to have fluctuated about a constant mean—no secular decrease in marine inundation of the continents can be identified. Such a decrease was first proposed by Egyed[1], who tried to show, with palaeogeographical measurements, that the seas have withdrawn steadily during the Phanerozoic; he deduced on this basis that the Earth is expanding. The idea of an expanding Earth has had little support[2], but the steady marine recession seems to be more widely accepted. Hallam[3], for example, confirmed Egyed's finding, and explained it as arising from orogenesis. My new estimate of the extent of the dry land, however, casts doubt on these earlier results, largely by taking into account continental drift and the progressive extension of shelf seas around the young Atlantic and Indian Oceans.

Wise[4] has argued that the secular withdrawal arises from the limitations inherent in palaeogeographical maps. An inevitable loss of resolution at more remote epochs imposes a temporal bias of uncertain direction, and a spatial bias towards shorter, simpler shorelines. These biases can be eliminated by sampling uniformly long intervals and by simplifying more recent shorelines; the resulting estimates are uniformly very inaccurate. Furthermore, the spectrum of sea-level change has a marked peak corresponding to periods of ~1–2 Myr (ref. 5). Such periods cannot be resolved easily (and perhaps not at all) by palaeogeographical methods, and they are much shorter than the present limits of palaeomagnetic resolution. Thus any palaeogeographical map will probably include information from more than one transgressive cycle, and in particular from the possibly asynchronous transgressive maxima of different regions. An overestimate of the extent of the seas will appear in any map which samples a period longer than ~1–2 Myr, and if the sample period grows shorter with time the extent of the seas will apparently decrease with time.

Notwithstanding these observations, Hallam's[3] data suggest secular withdrawal of the seas from the USSR and North America, and he concludes that the trend is real, and due most probably to orogenic thickening of the continents combined with variation in the length of oceanic ridges. It is not clear whether there ought to be a trend in ridge length (although large fluctuations are possible), but there may be a secular decrease in ridge volume, with a sea-level equivalent as great as −80 m (ref. 6) since −500 Myr if ridge volume is proportional to a 20% decrease in internal heat production; this latter estimate seems quite possible[7,8], although an 8–10% decrease may be more realistic. There are compelling arguments for a secular increase in the mass of the continents, but they are consistent with

0028-0836/81/190056—03$01.00

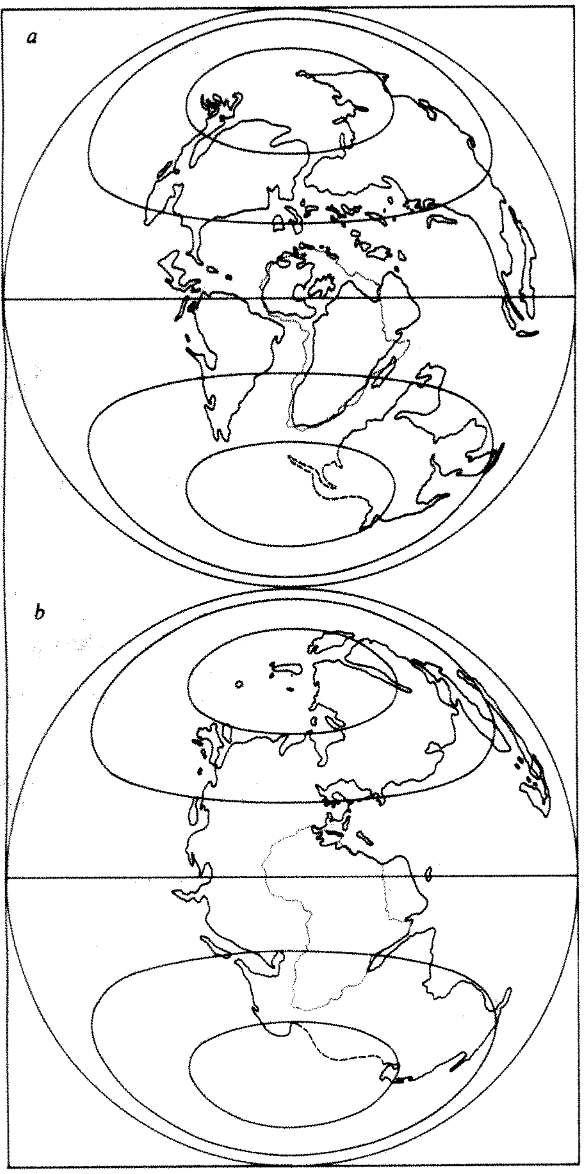


Fig. 1 Palaeographical reconstructions for *a*, −100 Myr and *b*, −200 Myr. Lambert equal-area projection of the whole sphere; parallels every 30°.

constant thickness (accommodating the added mass by an increase in area), and most of the increase may have been complete long before the Phanerozoic[9]. The continents may have thickened by net horizontal shrinkage without an increment of mass, as perhaps in the Appalachians[10], where wide, thick wedges of sediment and basement were thrust horizontally and stacked on top of each other in the Palaeozoic. Orogenic thickening, however, may be counterbalanced by the thinning recently described [11,12] as a precursor to the growth of sedimentary basins.

Fluctuations of sea level, as opposed to trends, may appear in history due to several causes[13–15]. Southam and Hay have shown that changes since −120 Myr can be explained by glaciation, episodic fast spreading at the ridges and a decrease in total ocean basin volume. Harrison *et al.* consider ocean sedimentation in more detail, and estimate the influence on apparent sea level of changes in the area–altitude distribution of the continents. Bond has also examined the latter influence, from a different perspective. Whatever the mechanisms that alter the sea level, I show here that the palaeogeographical evidence which prompted some of the earlier proposals is doubtful.

Table 1 Estimates of marine withdrawal from the continents

Sample	Withdrawal rate (10^3 km^2 Myr^{-1}) $-270 \leq T \leq 0$	$-582 \leq T \leq 0$
Cog	-6 ± 32	—
Bar	1 ± 43	—
Ter	67 ± 27	48 ± 9
TerH	67 ± 27	63 ± 9
Str	188 ± 106	101 ± 23
Hal1	55 ± 39	134 ± 18
Hal2	131 ± 100	119 ± 26
Wis	21 ± 29	15 ± 9

Sources: Cog, this work; Bar, ref. 30; Ter, ref. 1(*GPA*), Table 2; TerH, Ter with 12×10^6 km^2 subtracted from five earliest points[33]; Str, ref. 1 (*GPA*), Table 1; Hal1, ref. 3 (1977), Fig. 1 (USSR); Hal2, ref. 3 (1977), Fig. 2 (North America revised); Wis, ref. 4, Fig. 3 (North America: this work was an attempt to show that the rates of Str and TerH could be made to approach zero by sampling the record more densely).

There is no reason to suppose *a priori* that the continental shelves, which are below sea level today, have histories of inundation which differ significantly from those of areas which are above sea level today. It is not known whether at any instant the periphery of a continent is more likely to be inundated than its interior. Historically, however, a pattern of old fissures and sutures criss-crossing the continental interiors reveals that regions now far from oceans were once peripheral to old continents; conversely, all the modern Atlantic and Indian shelves were in continental interiors ~200 Myr ago. Earlier palaeogeographical analyses have ignored modern shelves for lack of offshore data, but consequently large areas of dry land have probably been left out. Indeed, if only present land areas are considered, then, unless the ancient and modern shorelines coincide everywhere, estimates of the extent of ancient seas must exceed the extent of modern seas. This property of the estimates is an artefact, and is neither logically nor geologically necessary.

Even without data from boreholes or seismic exploration, reasonable judgment can yield histories for modern shelves which are much simpler than those which would have to be contrived, often in conflict with palaeobiogeography, to ensure that they have always remained below sea level. Note that at −200 Myr in Fig. 1 the shelf off eastern North America is above sea level and well away from the shores of Pangaea; the alternative is a narrow seaway connecting Mexico with Scotland.

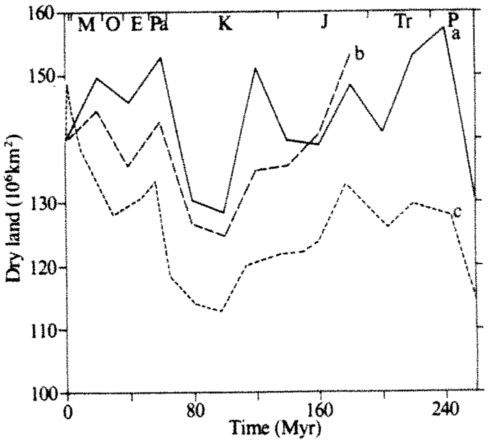


Fig. 2 Extent of dry land estimated from palaeogeographical maps: *a*, this work; *b*, ref. 30; *c*, ref. 1 after ref. 25 (Ter of Table 1). The points at −240 and −260 Myr in *a* have the same base map[34], dated at 250 ± 25 Myr; they are more uncertain, principally as to timing, than the others.

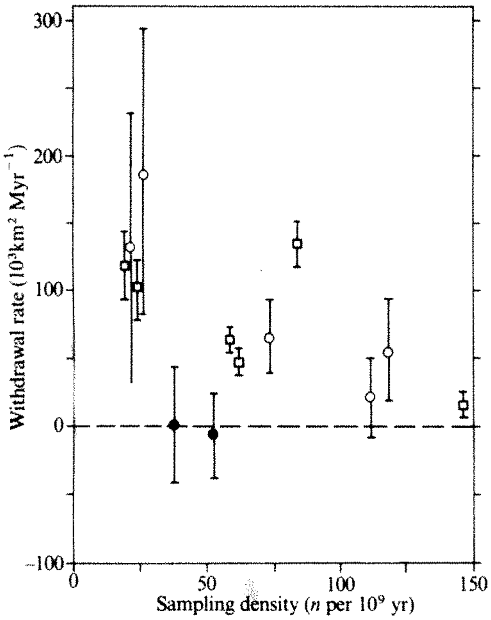


Fig. 3 Estimates of marine withdrawal from the continents. Symbols represent estimates of b (bounded by standard errors) from linear regressions $E = a + bT$, where T is time (adjusted from the Geological Society of London time scale and sources given in ref. 5) and E the extent of dry land (adjusted where necessary to global values). Modern E is set to 140.0×10^6 km^2 to simulate return of polar ice to the oceans[35]. (Rates are greater, though not significantly, if modern E is set to the observed 148.9×10^6 km^2.) ●, This work and ref. 30; ○, Table 1, column 1; □, Table 1, column 2 (note that the circles and squares are not from independent samples).

Without data from the shelf itself, the reconstruction as adopted is surely the more plausible. In fact, field data[16,17] show that the late Triassic is either absent, or represented by non-marine sediments penecontemporaneous with graben which suggest East Africa-style rifting. The narrow seaway can be ruled out. Field evidence also serves, with greater or lesser completeness, to rule out seaways between the various once-contiguous fragments of Gondwanaland[18–23].

Note too that the reconstructions of Fig. 1 allow for all known marine sediments within the present shoreline, and that they are conservative where data are lacking and no reasonable judgement can be made. There is, for example, no warrant for a landmass on the shelf south of the Falkland Islands at −100 Myr.

Figure 1 shows only specimens from a series constructed from 'palaeocontinental' maps[24] with shoreline information[25,26] superimposed. Offshore details were compiled from studies such as those in refs 16–23. These maps were planimetered to yield estimates of the extent of dry land at ~20-Myr intervals from 0 Myr to −260 Myr (Fig. 2a). The estimates are deficient in several ways. The continental reassemblies[24] are not unanimously accepted and are subject to errors of about 5–10° in positioning (95% confidence limits about inferred palaeopoles) and a few millions of years in time. The shoreline data were compiled 30 yr ago[25], and are from maps offset by up to 10 Myr from the palaeocontinental maps, but they do have the advantages of internal consistency and global completeness; more recent studies[27–29] tend to suggest revisions of detail, rather than of broad outline, to the work done by the Termiers. Offshore geology will always be less accessible than onshore, but its uncertainty, in the light of current knowledge, is no longer great enough to justify ignoring it. Percentage errors in offshore estimates can be appreciably worse than onshore errors and still contribute less to total error, because the offshore regions which might once have been land—up to $\sim 50 \times 10^6$ km^2—are less extensive than the onshore regions. About 10×10^6 km^2 of subglacial Antarctica, of which very little is known[26], are assumed here to have been above water since −260 Myr. It is difficult to place limits on errors from such diverse sources, but an estimate which is, at least partly, independent can be taken from recent work by Barron *et al.*[30] (Fig. 2b).

The trends (Table 1, Fig. 3) of Fig. 2a, b are statistically indistinguishable from each other and from zero. They should be compared with the trend seen in the curve of Egyed[1] (Fig. 2c), which was derived from ref. 25 but with data only from modern areas of dry land. The Egyed withdrawal rate and those below it in Table 1 all lack data from modern shelves. The rates in column 2 of Table 1 come from the longer span of the entire Phanerozoic, but only one of them significantly exceeds its counterpart in column 1. That is, most of the withdrawal rates are about as well defined in the past 270 Myr as in the past 580 Myr, indicating that it is reasonable to compare the rates which appear in column 1. Figure 2a lacks points in the Palaeozoic mainly because offshore palaeogeography becomes too speculative[31,32] for detailed mapping. Estimates of flooding within the present shoreline do not suffer this objection. However, as the arrangement of the continents in remoter times becomes less and less like the present one, the extrapolation from onshore results to putative global results becomes increasingly more doubtful. It may, in fact, be not only reasonable but prudent to concentrate on column 1. In any case, the trends of Fig. 2a, b are well below all the others of Table 1, and their zero secular trend appears also in graphs of relative change of Phanerozoic sea level[5]. (It is not certain, however, whether relative sea-level change and area of dry land are simply related[13].)

Thus, I conclude that the palaeogeographical evidence is not consistent with a secular withdrawal of the seas from the continents in the second half of the Phanerozoic. It is intriguing to speculate on the biases discussed by Wise[4], to which Fig. 2a is subject just as are other inundation curves. Does the zero trend of Fig. 2a imply that an 'ideal' version, that is, one with a uniform sampling interval ≤2 Myr, would show a secular encroachment of the seas?

I thank NSERC (Canada) for support.

Received 24 November 1980; accepted 18 February 1981.

1. Egyed, L. *Geofis. pura appl.* **33,** 42–48 (1956); *Nature* **178,** 534 (1956).
2. Glikson, A. Y. *Geology* **7,** 449–454 (1979).
3. Hallam, A. *Nature* **269,** 769–772 (1977); **232,** 180–182 (1971).
4. Wise, D. U. in *The Geology of Continental Margins* (eds Burk, C. A. & Drake, C. L.) 45–58 (Springer, New York, 1974).
5. Vail, P. R., Mitchum, R. M. Jr & Thompson, S. III *Am. Ass. petrol. Geol. Mem.* **26,** 83–97 (1977).
6. Armstrong, R. L. *Nature* **221,** 1042–1043 (1969).
7. Davies, G. F. *J. geophys. Res.* **85,** 2517–2530 (1980).
8. Turcotte, D. L. & Burke, K. *Earth planet. Sci. Lett.* **41,** 341–346 (1978).
9. O'Nions, R. K. & Pankhurst, R. J. *Earth planet. Sci. Lett.* **38,** 211–236 (1978).
10. Cook, F. A. *et al. Geology* **7,** 563–567 (1979).
11. McKenzie, D. *Earth planet. Sci. Lett* **40,** 25–32 (1978).
12. Sclater, J. G. & Christie, P. A. F. *J. geophys. Res.* **85,** 3711–3739 (1980).
13. Southam, J. R. & Hay, W. W. *J. geophys. Res.* **82,** 3825–3842 (1977).
14. Harrison, C. G. A. *et al. EOS* **61,** 1124 (1980).
15. Bond, G. *Geology* **6,** 247–250 (1976); *J. Geol.* **86,** 47–65 (1978).
16. Schlee, J. S. *Prof. Pap. P1167* (US Geological Survey, Reston, Virginia, 1980).
17. Sheridan, R. E. in *The Geology of Continental Margins* (eds Burk, C. A. & Drake, C. L.) 391–407 (Springer, New York, 1974).
18. Baker, B. & Miller, J. *Nature* **199,** 346–348 (1963).
19. Blant, G. (ed.) *Sedimentary Basins of the African Coasts, Pt 2, South and East Coasts* (Association of African Geology Surveys, Paris, 1973).
20. Hayes, D. E. (ed.) *Antarctic Oceanology II: The Australian-New Zealand Sector,* Antarctic Res. Ser. 19 (American Geophysical Union, Washington DC, 1972).
21. Nairn, A. E. M. & Stehli, F. G. (eds) *The Ocean Basins and Margins* Vol. 1 (Plenum, New York, 1973).
22. Theron, J. C. in *Gondwana Geology* (ed. Campbell, K. S. W.) 61–74 (Australian National University Press, Canberra, 1975).
23. Veevers, J. J. & Cotterill, D. *Bull. geol. Soc. Am.* **89,** 337–355 (1978).
24. Smith, A. G. & Briden, J. C. *Mesozoic and Cenozoic Palaeocontinental Maps* (Cambridge University Press, 1977).
25. Termier, H. & Termier, G. *Histoire Géologique de la Biosphère* (Masson, Paris, 1952).
26. Elliott, D. H. *Am. J. Sci.* **275A, 45–**[illegible]36 (1975).
27. Petters, S. W. *J. Geol.* **86,** 311–322 (1978).
28. Hallam, A. *J. Geol.* **79,** 129–157 (1971).
29. Harrington, H. J. *Bull. Am. Ass. petrol. Geol.* **46,** 1773–1814 (1962).
30. Barron, E. J., Sloan, J. L. II & Harrison, C. G. A. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **30,** 17–40 (1980).
31. Morel, P. & Irving, E. *J. Geol.* **86,** 535–561 (1978).
32. Ziegler, A. M. *et al. Tectonophysics* **40,** 13–51 (1977).
33. Holmes, A. *Principles of Physical Geology* 2nd edn, 968–970 (Nelson, London, 1965).
34. Briden, J. C., Drewry, G. E. & Smith, A. G. *J. Geol.* **82,** 555–574 (1974).
35. Nelson, A. B. *Rand Pap. P-5612* (Rand Corporation, Santa Monica, 1976).

Podzol chronosequences on coastal dunes of eastern Australia

SCIENCE COLLEGE
Upper Circular Road

C. H. Thompson

CSIRO Division of Soils, Chancellor's Place, St Lucia,
Queensland 4067, Australia

Chronosequences of freely drained podzols, formed in extensive dune systems of quartz sands along the east coast of Australia, offer unique opportunities for comparing temporal changes in different climatic zones during the late Quarternary. The largest areas and greatest number of dune systems occur along the subtropical coast near Brisbane. Here parts of six dune systems have not been covered by younger deposits and have been exposed to weathering since deposition. These areas constitute a relative age sequence on which podzols have formed with solum depths varying with age from <1 m to >20 m. Absolute ages for each system have not been determined but differences in erosion and weathering depths and comparisons with dated dunes elsewhere along the coast indicate an age sequence extending back at least to the last interglacial. Similarities in geomorphology and soils enable correlation with dune systems in other areas to the north and south and so provide a greater potential for examining temporal changes in soils and biological systems than that of previous soil chronosequences[1–5].

Because Australia largely escaped glaciation and has been relatively stable, tectonically, during the Quarternary, vast quantities of sand have accumulated along its coast[6]. Along the east coast siliceous sands dominate and have apparently been derived from the ocean floor during periods of marine transgression[6,7], the large accumulations representing intervals of marine and/or aeolian deposition when sea-level returned to similar positions in successive interglacial periods[8]. Two general types of deposits occur: low sandridges and swales parallel to the coast and widely distributed along its length, and multiple systems of transgressive parabolic dunes at 11 localities between 12 and 40° S latitude (Fig. 1). The parabolic dunes are generally oriented to the south-east onshore winds but their shapes reveal both the control of underlying topography, as in the case of nested dunes, and at exposed sites the influence of onshore winds from other quarters, which lead to multiple-lobe dunes.

Podzols and humus podzols have formed on both types of deposits[1,9–11] but the high transgressive dunes are more conducive to long-term soil formation because they lack watertables to considerable depth. On these porous, freely drained sands, podzol profile development reaches great depths and on the oldest dunes can exceed 20 m.

The largest area and greatest number of the large dune systems occur between 25 and 28° S. Here large areas of several sand-dune systems form North Stradbroke, Moreton and Fraser Islands and the mainland sandmasses at Cooloola and Peregian[10,12] (Fig. 1).

Studies of sand-dune systems and associated soils at Cooloola reveal eight periods of dune building. Sizeable areas of six of these dune systems have not been covered by successively younger deposits and have thus been subjected to sub-aerial weathering since deposition. The younger systems (1–4, Fig. 2) consist of parabolic dunes; the dunes of each system are clearly defined by their shape and distribution inland from the coast, the younger dunes being progressively smaller and deposited closer to the coast. The older systems (5 and 6, Fig. 2) extend further inland, have been much modified by erosion and have lost any recognizable dune orientation. They may have been more extensive in the past and parts of 6 may have been reexposed by erosion; however, the strong topographical difference between 5 and 6 makes this unlikely.

The sands are mostly quartz grains with usually <2% of heavy minerals (ilmenite, rutile, zircon and monazite) and some feldspar, but rarely any shell fragments. In all dune systems, the sands are well sorted and have modal grain sizes between 180 and 210 μm. However, the grains in the oldest system are more strongly weathered than those in the youngest (M. P. Green, personal communication), similar to those on nearby Fraser Island[13].

The unweathered sands are pale yellow-brown due to thin sesquioxidic coatings on the quartz grains. As weathering proceeds, the coatings on grains in the upper part of the profile are mobilized and together with humic compounds deposited lower in the profile to form the friable to firm, iron-humus B horizon (Bhir) of freely drained podzols. With increasing age the top of the Bhir is progressively remobilized and deposited deeper in the profile, the A and B horizons become thicker and more intensely developed, and there is a marked increase in thickness of A2 relative to A1 and B horizons[11]. In this way solum depth increases with age and is only limited by the thickness of freely drained dune sand and erosion rates since deposition. In the investigated sequences, depth of solum varies from <1 m on the youngest dunes to >20 m on the oldest (Fig. 2). All the sands except those of the youngest dune system are >30 m thick which ensures the pedological unity of the weathering column. This has been established by augering through the soil profiles to unweathered sand (systems 1–4) or into the B horizon (systems 5 and 6). If the trailing arms of parabolic dunes are intact, erosion losses are minimal on dune floors near the apex, low on the crests of trailing arms, and high on other slopes; depth of soil development varies accordingly.

Because both A1 and A2 horizons consist of clean quartz sand from which the sesquioxide coatings have been lost[14], the thickness of these horizons (depth to Bhir) may be used as a measure of relative weathering between analogous sites in different systems. Depth to Bhir is <1.2 m in systems 1–3 (Fig. 2), ~5 m in system 4 and commonly ~11 m in in system 5 and 15 m in system 6. However, both systems 5 and 6 have been

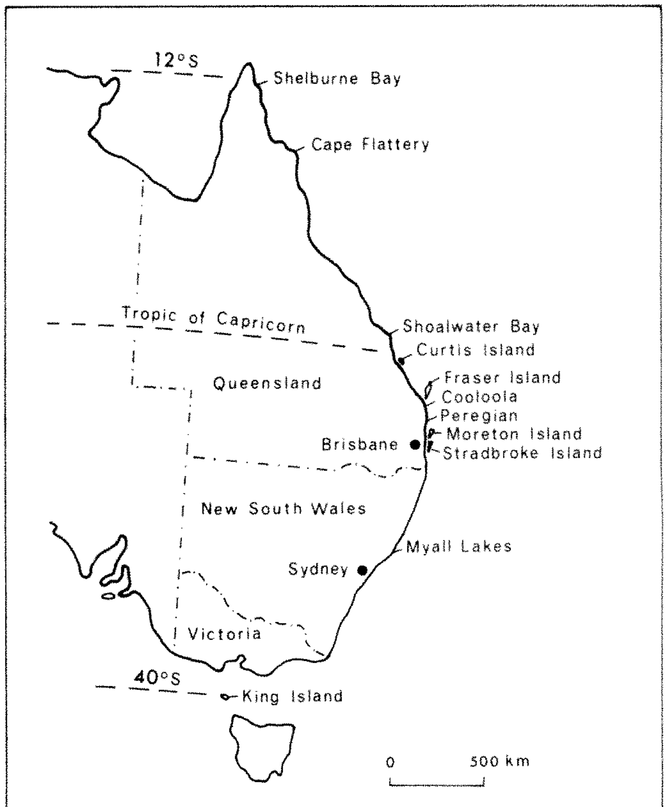


Fig. 1 Locations of main occurrences of parabolic sand-dune systems of quartz sands along the east coast of Australia.

0028–0836/81/190059—03$01.00

Table 1 Ridge crest profiles from each of the dune systems showing changes in organic carbon and total potassium and phosphorus with depth calculated for 1-m intervals

Depth (m)	Dune systems 6			5			4			3			2			1		
	OC	K	P	OC	K	P	OC	K	P	OC	K	P	OC	K	P	OC	K	P
1	0.27	15	10	0.41	20	10	0.39	90	10	0.47	890	30	0.24	1,380	20	0.35	1,170	60
2	0.02	10	10	0.02	10	10	0.05	60	10	0.15	1,160	30	0.06	1,460	40	0.09	1,200	45
3	0.02	10	10	0.02	10	10	0.04	190	10	0.04	1,170	20	0.05	1,360	40			
4	0.02	10	10	0.02	10	10	0.19	990	40									
5	0.02	10	10	0.02	10	10	0.21	1,350	55									
6	0.02	10	10	0.02	10	10	0.15	1,280	35									
7	0.02	10	10	0.02	10	10	0.11	1,250	35									
8	0.02	10	10	0.02	10	10	0.08	1,250	45									
9	0.02	10	10	0.02	10	10												
10	0.02	10	10	0.02	10	10												
11	0.02	10	10	0.02	10	10												
12	0.02	10	10	0.06	60	15												
13	0.02	10	10	0.20	300	30												
14	0.03	10	10	0.30	290	20												
15	0.04	20	20															

OC, organic carbon (%) by dry combustion in a Leco furnace. K and P, potassium (p.p.m.) and phosphorus (p.p.m.) by X-ray fluorescence. Limit of detection is ~10 p.p.m.

severely eroded and depth to B horizon varies accordingly, for example, in system 6 it ranges from 12 to >20 m.

Differences in soil morphology between the dune systems are reflected in the distribution of carbon, potassium and phosphorus down the profile (Table 1). The organic carbon profile indicates the development of A and B horizons. The potassium and phosphorus profiles reflect the weathering of feldspar and the removal of sesquioxide coatings from the quartz grains in the upper part of the profile. The unweathered sands contain virtually no apatite and phosphorus is held on the sesquioxide coatings.

Clay formation is limited by the very low content of weatherable minerals in the parent sands; the clay-size fraction in B horizons examined in systems 1–3 is <3%, in system 4 <7%, in system 5 <10% and up to 15% in system 6. Chlorite–vermiculite and kaolin dominate the clay fraction in the younger soils and kaolin and quartz in the older. The silt-size fraction is negligible (<1%) in the young systems 1–4, but may reach 7% in the top of the B horizon in systems 5 and 6, with negligible amounts above and below this position.

Stratigraphical and geomorphical relationships, and the differences in degrees of erosion, weathering and soil development show that the age sequence of the dunes at Cooloola extends over a long period. However, the absolute ages of these dune systems are unknown and no datable material has been recovered in the present study. A previous radiocarbon dating[12] of two pieces of wood embedded in the Cooloola sandmass near sea level indicated ages >40,000 and >45,000 yr BP from materials probably contaminated by modern carbon in groundwaters which issue at both sites. Additional dating of charcoal embedded in dune sands and a soil organic B horizon in overlying sands, both about 75 m above sea level, yielded ages of 39,000±3,000 and 30,300±800 yr BP[12]. Although the latter site cannot be located in relation to the dune systems now recognized it does show that the underlying sands are of Pleistocene age.

Dune systems 1–3 transgress, and in places fill, parts of deeply eroded valleys in the older dune systems. The preservation of dune outlines without erosion breaks and the limited depth of weathering in systems 1–3 (Table 1) also indicate young deposits. The diminishing size of these systems and of the individual dunes towards the coast suggests successive deposition from a diminishing sand supply during a high stillstand of the sea[7] as has occurred along the Australian coast during the Holocene[15]. Also, the degree of soil development on dunes in system 3 is similar to that observed in the dune at Triangle Cliff, Fraser Island, where a recent radiocarbon dating of charcoal at the base of the dune shows that it is <3,880 yr BP[16]. Therefore, systems 1, 2 and 3 are most likely of Holocene age.

In dune system 4, erosion has breached the trailing arms and reduced most of the dune floors and, together with the increased depth of soil weathering (Fig. 2; Table 1), indicates a sizeable time break between it and system 3. Even longer intervals are indicated by the increases in erosion and weathering between systems 4 and 5 and between 5 and 6. The increasing degree of modification indicates that these systems are older than Holocene and the early radiocarbon dates[12] show that their ages extend beyond 40,000 yr BP.

In New South Wales, two distinct barrier systems of low sand ridges have been recognized[9]—an Outer Barrier dated as a Holocene feature deposited over the past 6,000 yr (ref. 15) and an Inner Barrier which has been related, by uranium series dating of corals it overlies, to high sea levels of the last Pleistocene interglacial, ~120,000–135,000 yr BP[17]. While Inner and Outer Barrier equivalence has not been clearly established for the high dune systems along the Queensland coast, the younger

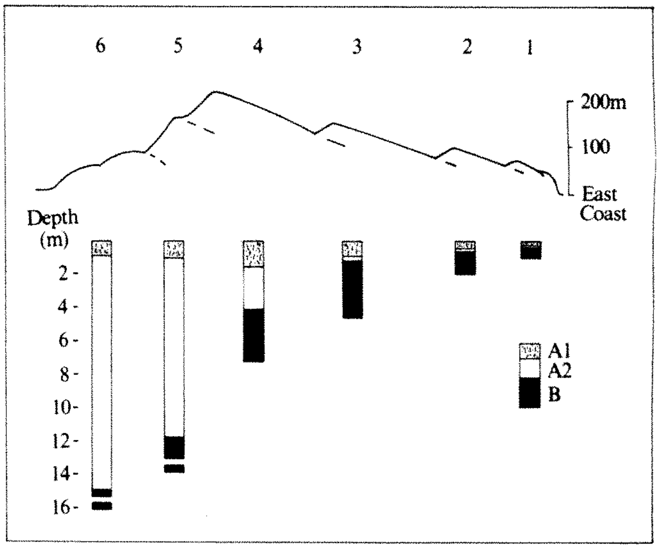


Fig. 2 Schematic cross-section of overlapping dune systems at Cooloola and podzol profiles associated with them. Numerals indicate dune systems, 1 being the youngest and 6 the oldest. The profiles represent soils at sites of least erosion for each system (dune floor in systems 1–4 and ridge crests in systems 5, 6). Thickness of B horizons in systems 5 and 6 is unknown.

systems (1–3) are probably equivalent to the Holocene transgressive parabolic dunes developed on parts of the Outer Barrier[9], judging from the limited amount of erosion and weathering evident at both localities.

Equivalence between the older dune systems at Cooloola and the Inner Barrier Deposits is less certain because of the greater areas of old deposits at Cooloola and the evidence of more events than can be reasonably represented by a single unit. For example, while no coastal sand deposition older than the Inner Barrier has been recognized in New South Wales[17], the coloured Teewah Sands are prominent in southern Queensland[12] and probably predate the Inner Barrier deposits. At Cooloola, the Teewah Sands underlie or form part of dune system 5 and at its northern end are underlain by older dune sands (system 6) which overlie a fossil beach. If system 6 deposits here equate with the Inner Barrier as suggested[18], there seem to be more depositional events at Cooloola between the last interglacial and Holocene than is suggested by records elsewhere. Alternatively, system 5 at Cooloola can be regarded as the most likely equivalent of the Inner Barrier and system 6 as an older deposit as has been suggested for the fossil beach[19].

In the Cooloola area, there is no firm indication of a change to markedly drier conditions during the last glaciation. The parabolic shape of the dunes in systems 1–4 and the lack of dune shapes characteristic of extensive bare sands suggest that this area has carried ground cover continuously since before the beginning of deposition of dune system 4. However, the evidence of extensive fluviatile erosion in systems 4–6 suggests that they have experienced wetter conditions at some time in the past than those that now prevail or are believed to have occurred since the onset of the last glaciation[20].

The soil landscapes here and those on Fraser, North Stradbroke and Moreton Islands have been correlated[10] and more recent field examinations indicate that sequences of podzols similar to that at Cooloola also occurs on these islands. Another such sequence is evident in the high dune systems further north at Shoalwater Bay (22.5° S) where at least four of the systems can be equated with those at Cooloola. To the south, soils on the parabolic dunes of the Outer Barrier[9] in the Myall Lakes area (33° S) are also similar to those of at least two of the systems at Cooloola. The soils of the parabolic dune systems further to the north and south have not yet been correlated; however, airphoto patterns of the dunes at Cape Flattery (15° S) and Shelburne Bay (12° S) show similar dune forms and descriptions from King Island[21] (40° S) suggest that one or more of the dune systems recognized at Cooloola are represented in these areas.

It is evident that the podzol sequence at Cooloola has developed over a considerable period including at least the last interglacial, last glaciation and postglacial periods, and podzol sequences on similar dune systems to the north and south can be correlated with it.

These podzol chronosequences provide unique opportunities for scientific study, both of changes that have occurred with time at one locality and of parallel changes that have occurred in temperate, subtropical and tropical areas. The sanddune systems at Cooloola are being intensively studied by the CSIRO Division of Soils, Brisbane, with collaboration from other institutions. Recognition of the dune systems and the podzol chronosequence has provided the framework for studies of erosion, nutrient movement, atmospheric accession, dune hydrology, and ecology of vegetation, fungi and soil fauna—these investigations need extending to the north and south.

Received 22 September 1980; accepted 2 March 1981.

1. Burgess, A. & Drover, D. P. *Aust. J. Bot.* **1,** 83–94 (1953).
2. Dickson, B. A. & Crocker, R. L. *J. Soil. Sci.* **4,** 123–154 (1953).
3. Franzmeier, D. P. & Whiteside, E. P. *Mich. Agr. Exp. Stat. Q. Bull.* **46,** 1–57 (1963).
4. Cowie, J. D. *N. Z. J. Geol. Geophys.* **6,** 268–280 (1963).
5. Jenny, H., Arkley, R. J. & Schultz, A. M. *Madrono* **20,** 60–74 (1969).
6. Bird, E. C. F. in *Barrier Islands* (ed. Schwartz, M. L.) 410–426 (1973).
7. Walker, H. J. *Soil. Sci.* **119,** 3–19 (1975).
8. Shakleton, M. J. & Opdyke, N. D. *Quat. Res.* **3,** 39 (1973).
9. Thom, B. G. *Proc. R. Soc. N.S.W.* **98,** 23–36 (1965).
10. Thompson, C. H. *Proc. R. Soc. Qd* **86,** 109–120 (1975).
11. Thompson, C. H. & Hubble, G. D. *Proc. int. Conf. CLAMATROPS,* Kuala Lumpur, 179–188 (1979).
12. Coaldrake, J. E. *Proc. R. Soc. Qd* **72,** 101–116 (1962).
13. Little, I. P., Armitage, T. M. & Gilkes, R. J. *Geoderma* **20,** 225–237 (1978).
14. Brewer, R. & Thompson, C. H. *Proc. int. Conf. CLAMATROPS,* Kuala Lumpur (1979).
15. Thom, B. G. & Chappell, J. *Search* **6,** 90–93 (1975).
16. Grimes, K. G. *Qd Govt Min. J.* **80,** 79–82 (1979).
17. Marshall, J. F. & Thom, B. J. *Nature* **263,** 120–121 (1976).
18. Langford-Smith, T. *Search* **3,** 102–104 (1972).
19. Ward, W. T., Little, I. P. & Thompson, C. H. *Paleogeogr. Paleoclimatol. Paleocol.* **26,** 305–316 (1979).
20. Bowler, J. M., Hope, G. S., Jennings, J. N., Singh, G. & Walker, D. *Quat. Res.* **6,** 359–394 (1976).
21. Jennings, J. N. *Geogr. J.* **123,** 59–70 (1957).

Climatically controlled variation of Sr and Mg in Quaternary planktonic foraminifera

Hans G. Cronblad & Björn A. Malmgren

Department of Geology, Stockholm University, Box 6801,
S-113 86 Stockholm, Sweden

Despite studies on the distribution of Sr and Mg in shells of marine calcareous invertebrates, the question of possible seawater temperature control on the concentrations of these elements in carbonate material remains controversial. Some workers have shown that individuals of a species living in warmer waters are richer in Sr and Mg than those living in cooler waters[1–7]. Others have found no such relationship[8–12]. We report here in tests of two planktonic foraminiferal species that the Sr concentrations are strongly related to late Quaternary climatic oscillations. Mg concentrations also show correlations with palaeoclimate, but not consistently. Concentrations of these trace elements, especially of Sr, in planktonic foraminifera may, therefore, be valuable for analysing palaeoclimatic history.

We have examined two species, *Globigerina bulloides* and *Globorotalia inflata*, in two deep-sea sediment cores from the southern Indian Ocean: E49–19 from the central subAntarctic and E48–22 from the Subtropical Convergence (Table 1). *G. bulloides* is most abundant in subpolar waters, whereas *G. inflata* predominates in the transitional zone between subpolar and subtropical waters[13]. *G. bulloides* lives in 'intermediate waters' (upper 100 m of the water column) and *G. inflata* inhabits 'deep waters' (generally below 100 m)[13].

Table 1 Data for Eltanin cores used in the study

Core	Lat. (S)	Long. (E)	Water depth (m)	Core length (cm)	Age of sequence (kyr)
E48–22	39°53.7′	85°24.6′	3,380	400	220
E49–19	43°53.2′	90°06.0′	3,036	550	300

Size-fractionation of samples may affect the oxygen isotope composition of planktonic foraminifera[14,15]. Tests of small specimens tend to be isotopically lighter than tests of the larger size of the same species. A pilot study has shown that there are no major differences between Sr and Mg concentrations in the various size fractions in *G. bulloides* and *G. inflata* (Table 2). Therefore, we concentrated on the >250-μm fraction of these relatively large species. More than 150 specimens of each species were picked per sample.

Specimens were ultrasonically cleaned to remove sediment and clay contaminants, and analysed using standard atomic absorption spectrophotometric methods. Only those tests that were apparently free of extraneous material were included. The analytical precision (95% confidence) is ±14 p.p.m. for Sr and ±11 p.p.m. for Mg. These estimates were based on analyses of four replicate subsamples (two per species) from each of four samples. Sr and Mg variations were compared with palaeoclimatic indices established from 'total fauna' analyses and

0028-0836/81/190061—04$01.00

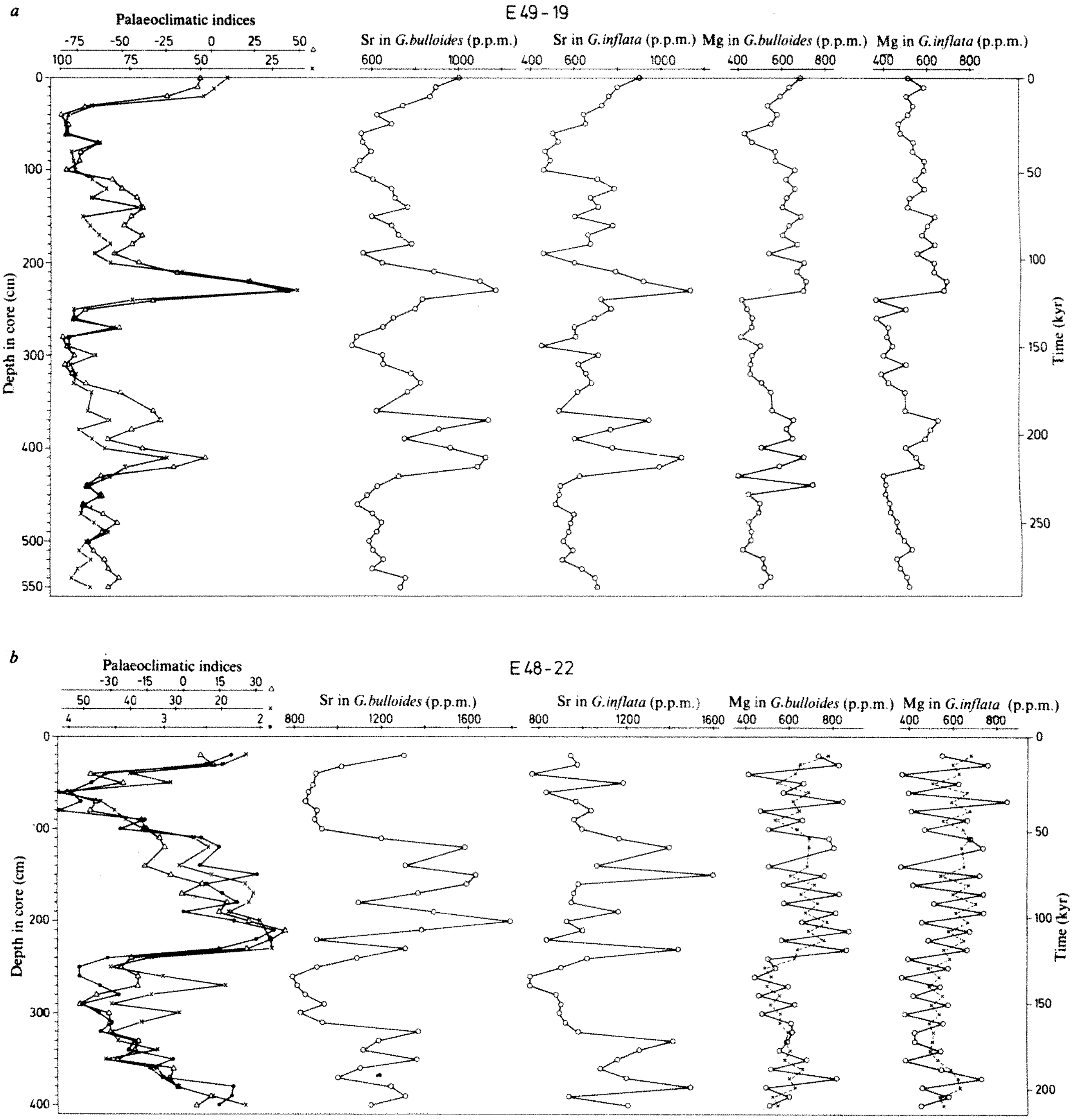


Fig. 1 *a*, Variation in Mg and Sr concentrations in *G. bulloides* in relation to two palaeoclimatic indices in core Eltanin 49–19 from the southern Indian Ocean. 'Warm' is to the right and 'cool' to the left. Chronology plotted on right is from Malmgren and Kennett[17]. *b*, Variation in Mg and Sr concentrations in *G. bulloides* and *G. inflata* in relation to three palaeoclimatic indices in core Eltanin 48–22 from the southern Indian Ocean. Dashed line curves for Mg show three-point moving averages. Chronology is from Williams[16]. △, 'Total fauna'; ×, % sinistral *N. pachyderma*; ●, $^{18}O/^{16}O$ ratios.

frequency in coiling proportions of *Neogloboquadrina pachyderma*, and also in E48-22 with a curve based on $^{18}O/^{16}O$ ratios in a planktonic foraminiferal species[16,17]. A 'total fauna' index is defined as the difference between relative abundances of a transitional and a subAntarctic assemblage of planktonic foraminifera[16].

In core E49–19, variations of the Sr concentrations in both species are remarkably similar in period and amplitude to the palaeoclimatic curves (Fig. 1). Mg concentrations are less variable and show less pronounced peaks, but the curves are clearly in phase with the palaeoclimatic oscillations. Cross-correlation analyses indicate that all curves have a very significant correlation (Table 3). Both elements are more enriched during warm climates (interglacial episodes) than during cool climates (glacial episodes).

Core E48-22 also shows good agreements between the Sr patterns in both species and the palaeoclimate (Fig. 2). Cross-correlation analyses between Sr and palaeoclimatic changes indicate that correlations in *G. bulloides* are highly significant (Table 3). Sr in *G. inflata* is correlated with $^{18}O/^{16}O$ ratios, but not with the other curves. The Mg concentrations in *G. bulloides* are also correlated with palaeoclimatic indices, whereas those in

Table 2 Sr and Mg compositions in *G. bulloides* and *G. inflata* from different size fractions

Size fraction (μm)	52 cm *G. bulloides* Sr (p.p.m.)	52 cm *G. bulloides* Mg (p.p.m.)	52 cm *G. inflata* Sr (p.p.m.)	52 cm *G. inflata* Mg (p.p.m.)	122 cm *G. bulloides* Sr (p.p.m.)	122 cm *G. bulloides* Mg (p.p.m.)	122 cm *G. inflata* Sr (p.p.m.)	122 cm *G. inflata* Mg (p.p.m.)
175–250	900	630	1,150	650	1,450	810	1,070	760
250–351	880	630	1,180	620	1,490	810	1,090	740
>351	890	660	1,200	640	—	—	—	—

Size fraction (μm)	222 cm *G. bulloides* Sr (p.p.m.)	222 cm *G. bulloides* Mg (p.p.m.)	222 cm *G. inflata* Sr (p.p.m.)	222 cm *G. inflata* Mg (p.p.m.)	362 cm *G. bulloides* Sr (p.p.m.)	362 cm *G. bulloides* Mg (p.p.m.)	362 cm *G. inflata* Sr (p.p.m.)	362 cm *G. inflata* Mg (p.p.m.)
125–175	850	580	840	510	—	—	—	—
175–250	880	570	—	—	1,090	520	1,050	560
250–351	850	570	870	510	1,070	540	1,110	590
>351	890	560	850	500	1,090	530	1,070	550

Specimens were randomly picked from two cooler-water samples (52 cm and 362 cm) and two warmer-water samples (122 cm and 222 cm) from E48–22.

G. inflata are not. Small-scale fluctuations in this element are greater than in E49–19, but plots of three-point moving averages (Fig. 2) bring out these palaeoclimatic signals in *G. bulloides* more clearly. The peculiar 'sawtooth' pattern is probably only an analytical artefact.

Various hypotheses have been suggested for the calcification mechanism in marine calcareous organisms[18–21]. The observed trends indicate that the biomineralization in planktonic forminifera is, at least partly, controlled by seawater temperature. It has been shown empirically[22] that substitution of exchangeable elements in the molecular lattice is greater at higher temperatures. Other environmental parameters may also play a part in the calcification process. The uptake of trace elements is also partly controlled by genetical differences among organisms. Differences in trace element concentrations among various planktonic foraminiferal species have been noted[9,10], and are apparent in our material (Fig. 1).

Sr and Mg concentrations in *G. bulloides* and *G. inflata* are lower on average in cores from cooler waters (E49–19), and still lower in samples from cool subAntarctic waters (compare tops of cores in Fig. 1 and core-top data in Fig. 2). This suggests that temperature-related gradients exist in Sr and Mg concentrations. If they did, the down-core variations may reflect climatically induced periodic expansions and contractions of the cooler water masses to the south of the core sites. Despite the possible existence of similar latitudinal gradients in both elements, palaeoclimatic correlations are weaker in Mg. This may be because the Sr ions are more tightly bound to the calcite lattice[23], but could also be explained by the stronger influence of clay contamination on the Mg composition[24,25].

Within individual species of planktonic foraminifera, the degree of dissolution on the sea floor has been shown[10] to affect the Mg concentration; a less dissolved sample is more enriched in Mg compared with a sample of the same species subjected to stronger dissolution. Differential susceptibility to dissolution between glacial and interglacial episodes could, therefore, also explain the trends in Mg. To determine dissolution effects on *G. bulloides* and *G. inflata*, we analysed recent subAntarctic core-top samples from the Maurice Ewing Bank (eastern Falkland Plateau), South Atlantic Ocean. No biogeographic

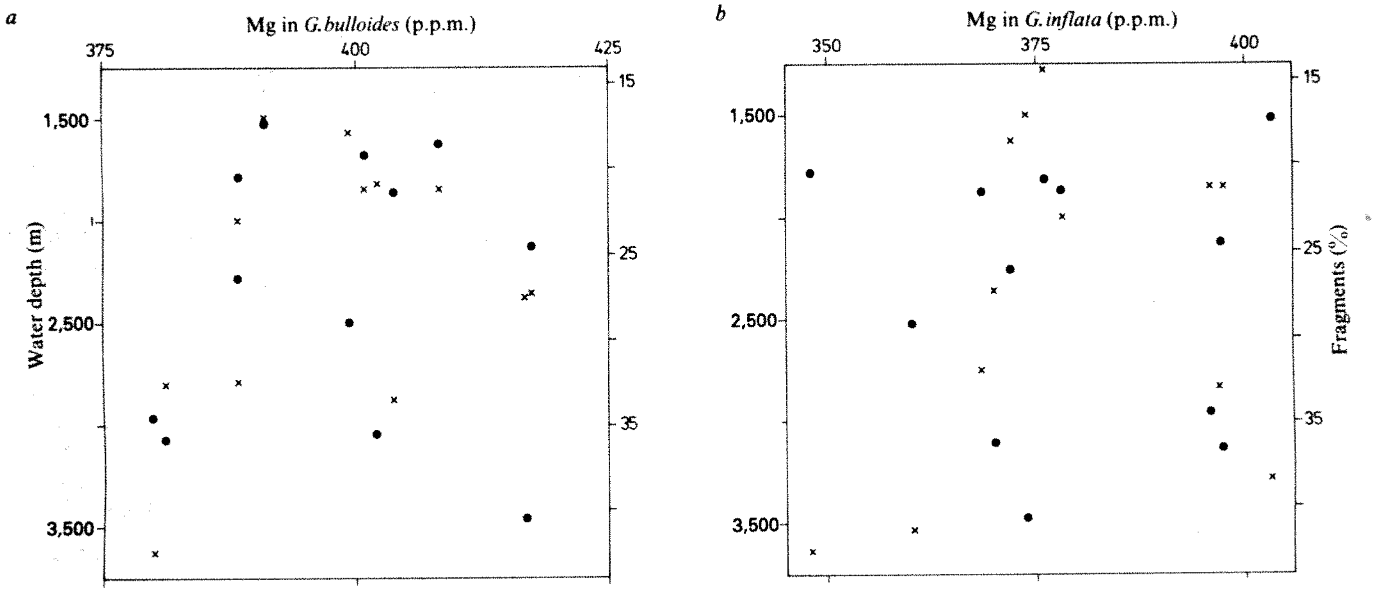


Fig. 2 Mg concentrations in *G. bulloides* (*a*) and *G. inflata* (*b*) in relation to water depth and degree of fragmentation of planktonic foraminiferal tests in cool subAntarctic core-top samples (Eltanin and Islas Orcadas) from the Maurice Ewing Bank, South Atlantic Ocean. Mg is not related to these parameters, which indicates that it is not affected by calcium carbonate dissolution as suggested previously[10]. ●, water depth; ×, % fragments.

Table 3 Cross-correlation coefficients among trace element compositions in *G. bulloides* and *G. inflata*, and palaeoclimatic indices in cores E49–19 and E48–22

E49–19

	Total fauna	% Sinistral *N. pachyderma*
Sr in *G. bulloides*	0.80*	−0.70*
Sr in *G. inflata*	0.77*	−0.69*
Mg in *G. bulloides*	0.64*	−0.51*
Mg in *G. inflata*	0.57*	−0.43*

E48–22

	Total fauna	% Sinistral *N. pachyderma*	$^{18}O/^{16}O$
Sr in *G. bulloides*	0.54*	−0.50*	−0.69*
Sr in *G. inflata*	0.19	−0.13	−0.38†
Mg in *G. bulloides*	0.43†	−0.36†	−0.44‡
Mg in *G. inflata*	0.30	−0.27	−0.29

These coefficients were determined for lags of zero, where correlations are normally at a maximum.
* $P \leq 0.001$; †$P \leq 0.05$; ‡$P \leq 0.01$.

differences exist over this narrow latitudinal belt (50–51°S), and differences in Mg should be due to dissolution. The Mg content does not change with water depth, and is also not related to the degree of fragmentation of the planktonic foraminiferal tests (Fig. 3). The degree of fragmentation has been shown[26,27] to represent a dissolution index in marine sediments. Therefore, it is less likely that the Mg signals in the cores are affected by dissolution. Likewise, no dissolution effects were found on Sr, which is in agreement with previous observations[28].

The results indicate that Sr concentrations in *G. bulloides* and *G. inflata* closely follow the late Quaternary climatic variations. The material available allowed only single analyses per sample; the precision of individual sample estimates could be improved by replicate analyses. Also, with further refinements of laboratory techniques, much of the 'noise' in the Sr curves might be reduced, which could aid in clarifying the palaeoclimatic signals. If so, Sr concentrations in these and other planktonic foraminiferal species, are a new independent means for palaeoclimatic analysis of Quaternary and, possibly, older sediments.

This research was supported by grant G4076-101 from the Swedish Natural Science Research Council. We thank Drs M. L. Bender, J. P. Kennett, F. T. Manheim and L. Sjöberg for helpful comments, and D. S. Cassidy for providing the material from the South Atlantic. Chemical analyses were made by Adnan Fahmy, Stockholm University.

Received 6 October 1980; accepted 23 February 1981.

1. Clarke, F. W. & Wheeler, W. C. *U.S. geol. Surv. Prof. Pap.* **124,** 1–62 (1922).
2. Chave, K. E. *J. Geol.* **62,** 266–283 (1954).
3. Chave, K. E. *J. Geol.* **62,** 587–599 (1954).
4. Blackmon, P. D. & Todd, R. *J. Paleont.* **33,** 1–15 (1959).
5. Dodd, J. R. *Geochim. cosmochim. Acta* **29,** 385–398 (1965).
6. Schopf, T. J. M. & Manheim, F. T. *J. Paleont.* **41,** 1197–1225 (1967).
7. Zolotarev, V. N. *Geochem. Int.* **11,** 347–352 (1975).
8. Krinsley, D. *Micropaleontology* **6,** 297–300 (1960).
9. Savin, S. M. & Douglas, R. G. *Bull. Geol. Soc. Am.* **84,** 2327–2342 (1973).
10. Bender, M. L., Lorens, R. B. & Williams, D. F. *Micropaleontology* **21,** 448–459 (1975).
11. Ragland, P. C., Pilkey, O. H. & Blackwelder, B. W. *Chem. Geol.* **25,** 123–134 (1979).
12. Windom, H. L. & Smith, R. G. *Southeastern Geol.* **13,** 275–281 (1971).
13. Bé, A. W. H. in *Oceanic Micropaleontology* Vol. 1 (ed. Ramsay, A. T. S.) 1–100 (Academic, London, 1977).
14. Emiliani, C. *Science* **173,** 1122–1124 (1971).
15. Berger, W. H., Killingley, J. S. & Vincent, E. *Oceanol. Acta* **1,** 203–216 (1978).
16. Williams, D. F. *Mar. Micropaleont.* **1,** 363–375 (1976).
17. Malmgren, B. A. & Kennett, J. P. *J. Paleont.* **52,** 1195–1207 (1978).
18. Towe, K. M. & Lowenstam, H. A. *J. ultrastruct. Res.* **17,** 1–13 (1967).
19. Travis, D. F. *Ann. N.Y. Acad. Sci.* **109,** 177–245 (1963).
20. Harrison, P. M. in *Iron Metabolism—An International Symposium*, 40–56 (Springer, Berlin, 1964).
21. Reimann, B. E. F. *Cell. Res.* **34,** 605–608 (1965).
22. Krauskopf, K. *Introduction to Geochemistry* (McGraw-Hill, New York, 1967).
23. Curtis, C. D. & Krinsley, D. *Geochim. cosmochim. Acta* **29,** 71–84 (1965).
24. Kilbourne, R. T. & Sen Gupta, B. K. *Geol. Soc. Am., Southeastern Sect., 22nd A. Meet. (Abstr.)* **5,** 408–409 (1973).
25. Jørgensen, N. O. *Bull. geol. Soc. Denmark* **24,** 299–325 (1975).
26. Thiede, J. *Bull. geol. Soc. Am.* **84,** 2749–2754 (1973).
27. Thunell, R. C. *Quat. Res.* **6,** 281–297 (1976).
28. Lorens, R. B., Williams, D. F. & Bender, M. L. *J. sedm. Petrol.* **47,** 1602–1609 (1977).

Intra-sex and inter-sex sibling interactions as sex ratio determinants

P. D. Taylor

Department of Mathematics and Statistics, Queen's University,
Kingston, Ontario, Canada K7L 3N6

Fisher[1] produced the first general argument that a random-mating sexually-reproducing population should devote equal reproductive resources to the production of male and female offspring, while Hamilton[2] considered several unusual situations in which the expenditure of reproductive resources should be biased towards one sex or the other. His 'host' model has been widely cited[3–6] and at least three factors proposed as the cause of the sex ratio bias: local mate competition, sib-mating and inbreeding. The problem of which of these factors is the real cause of the biased sex ratio seems important in view of the large number of apparently different but often closely related sex ratio models which are now being proposed. It would be useful to have ways of classifying them according to the presence of different factors which affect the equilibrium sex ratio. Here, I present a general equation (equation (2)) for differential change in fitness which exhibits the equilibrium sex ratio as achieving a balance between the factor contained in Fisher's standard grandchild argument[1] and a second factor resulting from intra-sex sibling interactions over reproductive resources. In particular, I analyse Hamilton's host model[2] for the presence of these different factors.

I will present the model for the special case of discrete non-overlapping generations. Although many of the phenomena I wish to study lie outside this special case, the formulation of the model is complicated by the mixing of generations, and it seems best to start with a simple situation. We divide the life cycle of the organism into two phases. Phase 1 encompasses the time from birth to the point at which some significant sex-specific behaviour begins, and phase 2 continues from there to the birth of the next generation. The purpose of the model is to relate the effects of sex-specific sib interactions in phase 2 to the equilibrium sex ratio.

Suppose in the population a typical mated female produces a total of x^* daughters and y^* sons at the end of phase 1. Consider a rare female who manages to produce x daughters and y sons. Her fitness in the $*$-population will be

$$w(x, y) = D(x, y)x + S(x, y)y \qquad (1)$$

where D is her expected ultimate genetic contribution through daughters per daughter and S is her expected contribution through sons per son. If at the end of phase 1, the offspring disperse at random into the general population, with no further non random sib interaction, then $D(x, y)$ and $S(x, y)$ are constants equal to their values D^* and S^* at (x^*, y^*); but otherwise they may not be constant.

Constraints operating during phase 1 will restrict the possible values of x and y in the population to a one-parameter family parameterized by the proportion r of reproductive resources allocated by the mother to the production of sons. In the simple case of a linear interchange between sons and daughters, with both sexes equally expensive, we have $x = K(1-r)$ and $y = Kr$, where K is the litter size. With this parameterization, w, D and S become functions of r and the value of $\mathrm{d}w/\mathrm{d}r$ at $r = r^*$ (the value giving x^* and y^*) is a measure of the evolutionary pressure for an increase in r^*. If $\mathrm{d}w/\mathrm{d}r > 0$ we expect the population to evolve towards an increase in the proportion of resources allocated to the production of sons; if $\mathrm{d}w/\mathrm{d}r < 0$, this proportion should decrease. Now, differentiating equation (1),

0028-0836/81/190064—03$01.00

$$\frac{dw}{dr} = \left[D^*\frac{dx}{dr} + S^*\frac{dy}{dr}\right] + \left[\frac{\partial D}{\partial x}\frac{dx}{dr} + \frac{\partial D}{\partial y}\frac{dy}{dr} + \frac{\partial S}{\partial x}\frac{dx}{dr} + \frac{\partial S}{\partial y}\frac{dy}{dr}\right] \qquad (2)$$

where everything is evaluated at r^*. The second term on the right measures the effect of sibling interactions in phase 2 on the mother's reproductive success, and it vanishes in the absence of these. Let us interpret the four summands of this term.

The partial derivative $\partial D/\partial x$ measures the change in the quantity 'fitness through daughters per daughter' per extra daughter, with number of sons, y, held constant. Thus, it measures the influence of extra daughters on the reproductive success of each daughter. For example, competition among daughters for reproductive resources during phase 2 would tend to make $\partial D/\partial x < 0$. Similarly, the analogous term $\partial S/\partial y$ measures absence of competition among brothers for reproductive resources. The cross term $\partial D/\partial y$ measures the influence of extra sons on the reproductive success of each daughter and similarly $\partial S/\partial x$ measures the influence of extra daughters on the reproductive success of each son.

Let us consider some examples. Clark[7] has studied African galagos in which young males disperse, but females remain within their home ranges and compete with their mother and sisters for reproductive resources (space, food). The resulting competition between sisters may contribute negatively to $\partial D/\partial x$. In MacDonald's foxes[8], females form cooperative breeding groups in which non-breeding vixens seem to confer benefits on the mother. There is strong evidence that members of a group are closely related, contributing positively to $\partial D/\partial x$.

Where males enhance the reproductive success of their sisters, we may get a positive contribution to $\partial D/\partial y$. Woolfenden[9] has shown that male helpers at the Florida scrub jay nest can have a significant effect on reproductive output. In all three of the above examples much of the interaction is between generations and may have important effects on the equilibrium sex ratio. However, the simple assumption of our model of non-overlapping generations restricts our attention to interaction within the same generation.

If competition between males for mates is more intense between brothers than between unrelated males, this contributes negatively to $\partial S/\partial y$. Partial sib-mating often has this effect, because brothers are competing for their sisters. Positive contributions to $\partial S/\partial y$ may result from cooperative behaviour among brothers. This is apparently the case, for example, for the lions studied by Bertram[10], who often hunt together and protect one another.

Partial sib mating also contributes positively to $\partial S/\partial x$, since extra daughters will increase the reproductive success of the sons. It is not so easy to find other examples in which daughters affect the reproductive success of their brothers. Pickering[11] describes a parasitoid wasp in which larval sibs growing on the host pupa compete for food. There is evidence that female larvae behave less altruistically towards their brothers than their sisters. An extra sister will affect both male and female reproductive success, but the effect on males may be greater.

To make specific calculations, we must find expressions for D and S in terms of x, y, x^* and y^*. In the above example of Pickering, the organisms are haplodiploid, and the resulting genetic asymmetry between males and females makes it difficult to find comparable measures for D and S. In this case special arguments are required[6,12]. The calculation of equilibrium sex ratio is further complicated by a conflict between mother and daughter over sex ratio[3,13]. In our calculations I will assume a diploid population. In this case, D and S can be calculated using the common measure of mated females; each mated daughter counts 1 and each female mated exclusively by a son counts 1 (with shared matings discounted by the appropriate factor). From this method of counting it follows that $D^*x^* = S^*y^*$.

The equilibrium value of r^* is obtained by setting dw/dr, given by equation (2), equal to zero. An examination of the various terms of equation (2) at equilibrium will display the different forces (on the value of r^*) which are held in balance at equilibrium.

As a normative example, consider Fisher's standard situation[1] in which offspring disperse after phase 1 and mate at random. Then $D(x, y) = D^*$, $S(x, y) = S^* = D^*x^*/y^*$, $x = K(1-r)$ and $y = Kr$. Therefore, all partial derivatives of D and S vanish, and equation (2) becomes

$$-D^*K + D^*K(1-r^*)/r^*$$

which, when set equal to zero, gives the equilibrium value $r^* = \frac{1}{2}$. We will refer to the first bracket in equation (2) as the Fisher term, as this is the term which must vanish at equilibrium in his standard model.

In Hamilton's model[2], mated females collect on hosts in random groups of size N to produce and raise K offspring each. At the start of phase 2, the KN offspring mate at random among themselves, and the mated females go out into the general population to form new host groups and start the cycle again. Because a mutant female is rare we can assume the other females in her host group are normal. Hence, $D(x, y) = D^*$ and $S(x, y) = D^*(x + (N-1)x^*)/(y + (N-1)y^*)$. With these expressions both partial derivatives of D vanish, $\partial S/\partial x$ is positive because of the sib-mating and $\partial S/\partial y$ is negative because of male reproductive competition resulting from the sib-mating. Then equation (2) has the form

$$D^*\left[\frac{dx}{dr} + \frac{x^*}{y^*}\frac{dy}{dr}\right] + \left[\frac{\partial S}{\partial x}\frac{dx}{dr} + \frac{\partial S}{\partial y}\frac{dy}{dr}\right]$$

Since $dx/dr < 0$ and $dy/dr > 0$ both terms in the second bracket are negative, and both act in the same way on r^*. At equilibrium the Fisher term must be positive, requiring a female-biased sex-ratio. With $x = K(1-r)$ and $y = Kr$, the equilibrium ratio turns out to be $r^* = (N-1)/2N$. Thus, the female bias, which is often loosely thought to result from sib-mating, actually results from two different forces: the effect of extra daughters on male mating success, and the effect of fewer sons on male reproductive competition.

Finally, the general model applies equally well to the model of Charnov[14] for barnacles and Fischer[15] for coral reef fish, of sex allocation in hermaphrodites. (Replace mated female by individual, daughter and son by egg and sperm, and the process of mating by that of fertilization.) Each individual delivers its sperm to fertilize the eggs of N other individuals, and has its own eggs fertilized by, on average, N other individuals. As before a linear tradeoff is assumed between production of eggs x and sperm y. Thus $D(x, y) = D^*$ and $S(x, y) = D^*Nx^*/(y + (N-1)y^*)$. This is formally the same as Hamiltonian's model, except that we have removed the possibility of sib-mating, and $\partial S/\partial x$ is now zero. However, $\partial S/\partial y$ is the same as before, again forcing the Fisher term to be positive, creating a female-biased equilibrium ratio, this time, only through the effect of fewer sons on male reproductive competition.

The significance of equation (2) is that it resolves the evolutionary pressure on r^* into a sum of two factors which are held in balance at equilibrium: the standard Fisher factor and a second factor containing the effect of intra- and inter-sex sibling interactions. This second factor again resolves into four terms. Different sex ratio models can be compared according to the presence of these factors. Sib-mating (or selfing in hermaphrodites) or more generally mating between relatives, often contributes to these factors, but the inbreeding which results, in itself, has no role in determining sex ratio.

This research was supported by a grant from the Natural Sciences and Engineering Research Council of Canada. I thank Michael Bulmer for many discussions which sharpened my understanding of the causes of sex ratio bias and George Williams for valuable comments on the first draft.

Received 18 August 1980; accepted 26 February 1981.

1. Fisher, R. A. *The Genetical Theory of Natural Selection* (Oxford University Press, 1929).
2. Hamilton, W. D. *Science* **156,** 477–488 (1967).
3. Trivers, R. L. & Hare, H. *Science* **191,** 249–263 (1976).
4. Stenseth, N. C. *Oikos* **30,** 83–89 (1978).
5. Maynard Smith, J. *The Evolution of Sex* (Cambridge University Press, 1978).
6. Taylor, P. D. & Bulmer, M. G. *J. theor. Biol.* **86,** 409–419 (1980).
7. Clark, A. B. *Science* **201,** 163–165 (1978).
8. MacDonald, D. W. *Nature* **282,** 69–71 (1979).
9. Emlen, S. T. in *Behavioural Ecology* (eds Krebs, J. R. & Davies, N. B.) 245–281 (Blackwell, Oxford, 1978).
10. Bertram, B. C. R. in *Behavioural Ecology* (eds Krebs, J. R. & Davies, N. B.) 64–96 (Blackwell, Oxford 1978).
11. Pickering, J. *Nature* **283,** 291–292 (1980).
12. Benford, F. A. *J. theor. Biol.* **72,** 701–727 (1978).
13. Bulmer, M. G. & Taylor, P. D. *Heredity* (in the press).
14. Charnov, E. L. *Mar. Biol. Lett.* **1,** 269–272 (1980).
15. Fischer, E. A. *Am. Nat.* **17,** 64–82 (1981).

Retinal projection in a non-visual area after bilateral tectal ablation in goldfish

S. C. Sharma

Department of Ophthalmology, New York Medical College,
Valhalla, New York 10595, USA

If, in the adult goldfish, one optic tectum is ablated, the regenerating optic axons from the contralateral retina innervate the remaining tectum, where they form a retinotopically ordered map[1–3]. The pathway for this induced ipsilateral projection[1] coincides with many of the pathways which normally connect the two tecta, but early in regeneration the optic fibres also enter non-visual centres to which there are degenerating tectal efferent pathways to follow[4]. We have therefore now investigated the fate of regenerating optic axons in goldfish from which both optic tecta have been removed; they are found to innervate non-visual centres, where again they generate a retinotopic map.

Fifteen adult goldfish (10–12 cm long and 2.5 yr old at the time of surgery) were anaesthetized with 0.5% MS222. The tectum was ablated bilaterally and extreme care taken not to damage the diencephalon, which joins the tectum at its rostral pole. Animals were allowed to survive between 380 and 620 days after tectal removal before electrophysiological mappings of the retinal projections were made. In five animals, the right eye was injected with 10 μl of [3,4-^{3}H]proline (specific activity 50 Ci mmol^{-1}) 1 day before mapping; the brains of these animals were processed for autoradiography. In five other animals, after visuotopic mapping, the right optic nerve was cut and 50% horseradish peroxidase (HRP) solution (Sigma type VI) applied to the cut end. The retinal projection of the remaining five animals was studied anatomically, using HRP orthograde transportation after injection of 5–7 μl of 30% HRP intraocularly. In the HRP studies, animals were killed 1–3 days later and processed for HRP reaction[5].

Autoradiographic analysis showed that, in addition to a high concentration of termination sites in the primary diencephalic visual centres, the optic fibres terminated in most tectal efferent areas[6,7]. In normal fish these areas do not receive direct retinal fibre terminations[8,9] and include the tori semicircularis, nucleus isthmi, nucleus of the rostral mesencephalic tegmentum, dorsolateral tegmental nucleus and nucleus reticularis superior of the hindbrain. In addition, radioactive grains were seen in the rostral valvula cerebelli.

Orthograde transport of HRP studied in 10 animals basically confirmed the autoradiographic identification of target centres of regenerating optic axons (Fig. 1). Most optic fibre fascicles in the regenerated optic nerve appear linearly arranged. The optic fibre fascicles separate from the ventral part of the optic tract at the diencephalic level and travel caudally before arborizing. The first bundle of optic fibres detached itself from the nerve at its rostral ventral position then the next dorsal bundle did the same. The uppermost portion of the nerve travelled farthest and apparently terminated in the valvula cerebelli (Fig. 1*b*, *c*). This pattern of axonal trajectories from the optic nerve is similar to that observed in normal goldfish (S.C.S., unpublished observation).

The visuotopic projection of either the right or left eye was determined in 10 animals by electrophysiologically mapping the area directly beneath the ablated tectum. This area includes the valvula cerebelli medially and the torus semicircularis laterally and ventrally. The visuotopic projection represented in Fig. 2 was essentially similar to that seen in eight other animals and was confined to the rostro-medial portion of the contralateral valvula cerebelli. The projection map in the remaining animal was incomplete. The representation of the superior contralateral visual field was similar in its extent to a normal visuotopic projection onto the contralateral tectum. The receptive field sizes of the multi-units were in the 10–15° range, which is similar to the size of normal retinotectal units.

In three animals, in addition to the projection to the valvula cerebelli, large receptive fields (60–80°) were recorded from the torus semicircularis and were mostly from the inferior visual field. It was not possible to record any meaningful visuotopic maps from the torus. Large receptive field units were also recorded from the the deeper tegmental areas.

These results imply that, in the absence of both optic tecta, the optic axons can regenerate to non-visual centres and generate a retinotopic projection which approximates to the normal map. As the retinotopic map represents presynaptic optic fibre terminations, it is not certain whether fibres formed synaptic connections with cells in anomalous areas. Anomalous connections to non-visual targets following one tectal removal have been reported in various vertebrates[10–13].

The valvula cerebelli extend rostrally into the optic ventricle and lie apposed to the ventral surface of the dorsal tectum

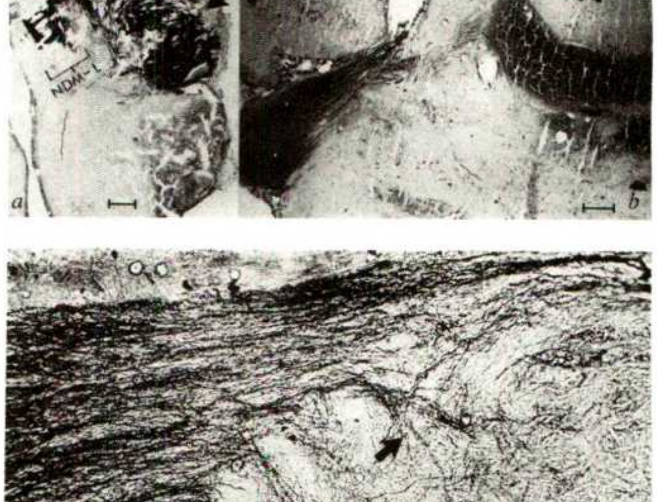


Fig. 1 *a*, Coronal section through the diencephalon showing contralateral retinal projection in a fish 380 days after tectal ablation where the eye was injected with HRP. The medial diencephalic retinal target areas (nucleus dorsalis medialis and lateralis, NDM, NDL are heavily labelled. The arrow points to a neuroma (a misleading entity as neuroma does not appear in sagittally sectioned brains). Scale bar, 100 μm. *b*, Sagittal section through the brain of a fish 390 days after tectal ablation. The arrow points to the fibre entering the valvula cerebelli (VC). Scale bar, 100 μm. *c*, High power photomicrograph of the optic nerve from *b*. The arrows point to optic fascicles separating from the nerve. Scale bar, 100 μm.

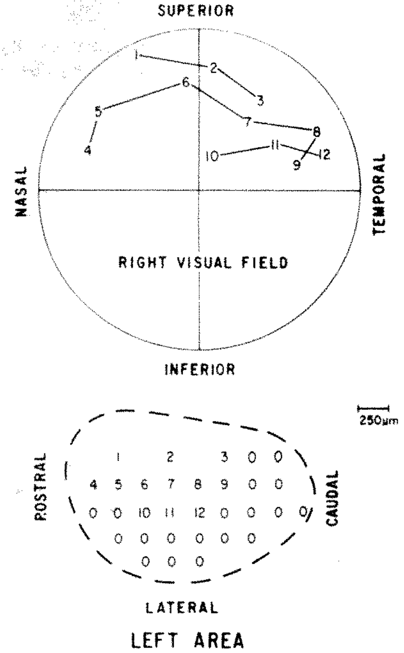


Fig. 2 The visual projection map from the right eye to the left portion of the valvula cerebelli, marked in the diagram as left area. Each number in the visual field shows the position of the stimulus that evoked maximum potential, recorded by an electrode at the position indicated by the same number on the left area. Empty circles represent electrode positions from which visually evoked potentials could not be recorded.

at the level of the diencephalic tegmental junction. Similarly, the rostral extent of the torus semicircularis begins at the level of the posterior commissure and is apposed to the ventral surface of the ventrolateral tectum. Removal of optic tecta is followed by injury and deafferentation to these structures, with the regenerating optic fibres entering them at their rostral poles. It is suggested here that innervation of the regenerating optic axons is facilitated by degenerating debris.

These results contradict the recent findings in frogs of Bohn and Stelzner[14], who, after unilateral tectal ablation and lesion in the isthmal region, showed that regenerating optic axons did not innervate the non-visual areas deafferented by isthmal transection. Note that in Bohn and Stelzner's study, one optic tectum was always left intact[14]; when a tectal tissue is left intact, optic fibres preferentially innervate it[15].

In normal goldfish, most optic axons terminate in the contralateral tectum. However, in the absence of both tecta, regenerating optic axons innervate non-visual areas and apparently form a topographically ordered map. Hence, it could be concluded that the tectum is not necessary for the formation of an ordered projection by the optic axons. The present results support the possibility that topography is intrinsic to the optic axons and that the formation of an ordered map may be guided by the presence of general axial cues (that is, rostro-caudal, medio-lateral) within the brain tissue.

This work was supported by NIH grant NEI01426. S.C.S. is the recipient of a Research Career Development Award from NIH-NEI.

Received 25 September 1980; accepted 2 February 1981.

1. Easter, S. S., Schmidt, J. T. & Leber, S. M. *J. Embryol. exp. Morph.* **45,** 145–159 (1978).
2. Levine, R. & Jacobson, M. *Expl Neurol.* **43,** 527–538 (1974).
3. Sharma, S. C. *Expl. Neurol.* **41,** 661–669 (1973).
4. Lo, R. Y. S. & Levine, R. L. *J. comp. Neurol.* **191,** 295–314 (1980).
5. Mesulam, M. M. *J. Histochem. Cytochem.* **24,** 1273–1280 (1976).
6. Grover, B. & Sharma, S. C. *J. comp. Neurol.* **184,** 435–454 (1979).
7. Grover, B. & Sharma, S. C. *J. comp. Neurol.* **196,** (in the press).
8. Reperant, J., Lemire, M., Miceli, D. & Peyrichoux, J. *Brain Res.* **118,** 123–131 (1976).
9. Sharma, S. C. *Brain Res.* **39,** 213–223 (1972).
10. Bassinger, J., Lund, R. D. & Miller, B. *Expl Neurol.* **54,** 369–382 (1977).
11. Kalil, R. E. & Schneider, G. E. *Brain Res.* **100,** 690–698 (1975).
12. Kicliter, E., Misentone, L. J. & Stelzner, D. J. *Brain Res.* **82,** 293–297 (1974).
13. Schneider, G. E. & Jhaveri, S. R. in *Plasticity and Recovery of Function in the Central Nervous System* (eds Stein, D. G., Rosen, J. J. & Butters, N.) 65–109 (Academic, New York, 1974).
14. Bohn, R. C. & Stelzner, D. J. *J. comp. Neurol.* **190,** 763–779 (1980).
15. Sharma, S. C. *Brain Res.* **183,** 453–457 (1980).

Birds can overcome the cardenolide defence of monarch butterflies in Mexico

Linda S. Fink* & Lincoln P. Brower†

* Webster Center for the Biological Sciences, Amherst College, Amherst, Massachusetts 01002, USA
† Department of Zoology, University of Florida, Gainesville, Florida 32611, USA

Flocks of black-backed orioles (*Icterus abeillei* Lesson) and black-headed grosbeaks (*Pheucticus melanocephalus* Swainson) eat several hundred thousand monarch butterflies (*Danaus plexippus* L.) in the dense overwintering colonies in central Mexico, and in 1979 were responsible for over 60% of the butterfly mortality at several sites[1]. Such predation is unusual because, during larval development the aposematically coloured monarch butterfly sequesters cardenolides from its milkweed foodplants (Asclepiadaceae)[2,3]. These bitter-tasting heart poisons cause vomiting in 12 species of birds in 9 families (refs 4–9, and L. P. Brower, unpublished observation), although the domestic chicken, Japanese quail, hedgehog, mouse and sheep[8,10,11] have been shown to be insensitive to their emetic effects. Extensive predation of monarch butterflies by birds has never been observed except in Mexico. We report here that the Mexican butterflies are weakly emetic, and that taste discrimination by orioles and cardenolide insensitivity of grosbeaks allow these birds to feed freely on monarch butterflies.

To verify that a substantial proportion of the Mexican butterflies contain cardenolides, we collected 198 of these in January 1980 from Site Alpha[1] and stored them on dry ice before spectrophotometric assay[12,13]. In common with wild populations from Massachusetts and California[14], they varied widely in total cardenolide content (Fig. 1).

Our first hypothesis was that the birds taste each butterfly captured, release the emetic (bitter) ones, and kill and eat only the individuals containing little or no cardenolide[15]. Observing the birds through binoculars, we saw orioles attack 98 butterflies and release 75 without apparent harm; grosbeaks similarly released 65 of 92. Because orioles and grosbeaks never eat the butterfly wings we were able to measure the cardenolide content of butterflies killed and partly eaten by the birds. We collected damaged butterflies from the forest floor during feeding bouts from 7 to 19 January 1980, and spectroassayed one fore and one hind wing of each. The cardenolide concentrations of the wings of 177 butterflies damaged by grosbeaks, 142 damaged by orioles and 150 undamaged (control) butterflies are shown in Fig. 2. The frequency distributions of the three samples are very similar which provides strong evidence that the birds kill monarch butterflies randomly with respect to cardenolide content.

Our second hypothesis, based on our observations that the birds eat only parts and drop large portions of the butterflies, was that the birds ate less of the more emetic individuals. Table 1 relates the cardenolide content of the 142 butterflies damaged by orioles to the degree to which they were eaten by the birds. Clearly orioles eat fewer of the butterflies with more cardenolide content ($P < 0.005$). Damage to thoraces showed the same relationship: mean cardenolide concentration of wings of butterflies with heavily damaged, slightly damaged and undamaged thoraces was 32, 44 and 55 μg per 0.1 g, respectively (one way analysis of variance, $F_{2,138} = 4.8$, $P < 0.01$). Thus, although individual butterflies were not initially protected from predation by containing more cardenolide, once killed, the amount of butterfly eaten by an oriole was inversely proportional to cardenolide content. The oriole further reduces its cardenolide intake by stripping out the thoracic muscle and abdominal contents without eating the cuticle in which these compounds are more highly concentrated[16]. Table 2 shows a similar analysis of the 177 butterflies damaged by grosbeaks which indicates no overall relationship between the extent of damage and cardenolide content ($P = 0.32$). In contrast to the orioles, grosbeaks consume cuticle as well as internal tissues. Thus, the orioles are selective once they have killed the butterflies, whereas the grosbeaks kill and continue to eat the butterflies randomly with respect to cardenolide content.

---

* Present address: Section of Ecology and Systematics, Cornell University, Ithaca, New York 14853, USA.
† To whom all correspondence should be addressed.

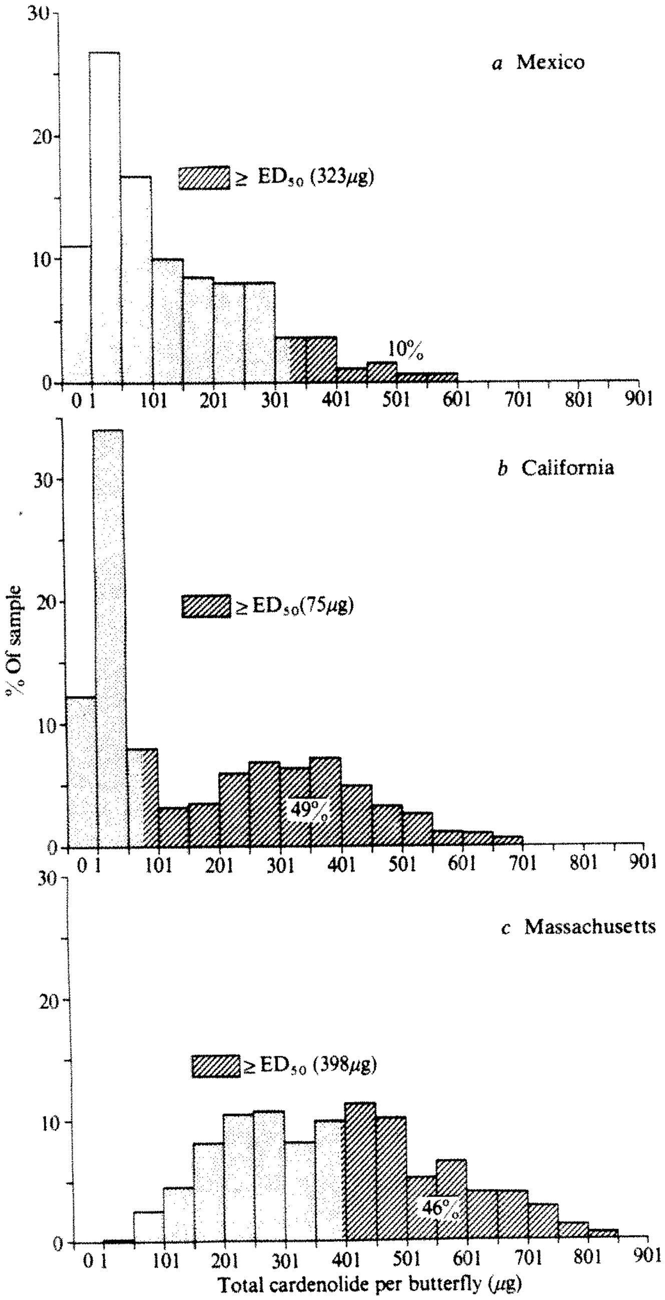


Fig. 1 *a–c* Compare total cardenolide per butterfly in overwintering monarch butterflies from Mexico (*a*) and California (*b*), and in autumn migrants from western Massachusetts (*c*). The hatched portions of the histograms are butterflies which contain sufficient cardenolide to cause a blue jay to vomit (that is, contain ≥1 blue jay ED$_{50}$ unit per butterfly). Because the Mexico population combines weakly emetic cardenolides with a high proportion of low cardenolide butterflies, only 10% carry more than one emetic dose, in contrast to 49 and 46% of the California and Massachusetts butterflies. The potential protection of the Mexico population against sensitive bird predators is therefore much reduced. The data for the Massachusetts and California butterflies are calculated according to the method of Brower *et al.*[13] from previously published absorbance values[14]. *a*, Site Alpha, 1980: $n = 198$, $\bar{x} = 128$, s.d. = 126, range 0–551. *b*, Muir Beach, 1971: $n = 351$, $\bar{x} = 165$, s.d. = 179, range 0–684. *c*, Hockanum, 1971: $n = 355$, $\bar{x} = 384$, s.d. = 172, range 34–850.

Table 1 Cardenolide concentrations (μg per 0.1 g dry weight of the wings) of monarch butterflies eaten to various degrees by black-backed orioles at the Site Alpha Mexican overwintering colony, January 1980

Damage category	% Of abdominal contents eaten	Sample size	Cardenolide concentration (μg per 0.1 g dry weight)		
			$\bar{x}$	s.d.	Range
a	91–100	29	20	30	0–106
b	51–90	43	31	41	0–156
c	11–50	32	49	50	0–204
d	1–10	25	57	58	0–234
*e**	0	13	65	72	0–208
Entire sample		142	41	50	0–234

The butterflies are grouped into five categories of decreasing abdominal damage. The data indicate a highly significant (one way ANOVA $F_{4,141} = 3.91$, $P < 0.005$) inverse relationship between the cardenolide content of the butterflies and the extent to which orioles eat them. Duncan's multiple range test indicates that categories *a* and *b* differ from categories *c–e* ($P < 0.05$).

* Abdomen unharmed, thorax partly eaten.

We next tested the sensitivity of the orioles and grosbeaks to the presumed emetic effect of the butterflies. For comparison with previous data, we first used our standard blue jay emetic dose 50 (ED$_{50}$) assay in which predetermined doses of dry powdered butterflies are force-fed in gelatin capsules to jays[3]. (The ED$_{50}$ is that dose at which there is a 50% probability that a bird will vomit.) In the first test we combined 24 butterfly abdomens from Site Alpha (January 1980 collection), and fed four jays increasing doses up to 590 mg of powder containing 236 μg of cardenolide—no bird vomited. This high dose, equivalent in weight to more than six abdomens, reached the upper physical limit of the test. To make a more potent force-feeding powder we screened butterflies for cardenolide content and selected those with substantial amounts. Two wings from each of 453 Site Alpha butterflies were tested qualitatively with the spectroassay reagents, and only the abdomens of those showing a positive reaction were used ($n = 158$). This screening increased the concentration of cardenolides in the force-feeding powder from 40 to 105 μg per 0.1 g. We assumed, as determined in the Massachusetts and California populations[14,17], that we did not differentially select for particular cardenolides. For 25 jays the ED$_{50}$ was 323 μg of cardenolide per 85 g (95% confidence limits = 247–425 μg).

We next tested the rosebreasted grosbeak (*Pheucticus ludovicianus* L.) which is closely related to the black-headed grosbeak[18]. Six birds were force-fed the powder previously given to the jays in doses from 132 to 331 μg of cardenolide (that is, from 0.41 to 1.0 blue jay emetic units). None of these birds vomited. Two additional grosbeaks were force-fed abdomens of butterflies reared on *Asclepias curassavica* L., previously shown[19] to contain cardenolides 7.5 times as emetic as those of the Mexican monarch butterflies (Jay ED$_{50}$ 43 μg per 85 g). One bird was fed an intact abdomen, estimated by spectroassay of a sibling to contain 270 μg of cardenolide. The second was fed a capsule of powdered abdomens containing 700 μg (16 blue jay emetic units) of cardenolide. Neither bird vomited, although both showed other signs of poisoning[5,20].

In January 1981 at Site Alpha, 10 black-headed grosbeaks and 11 black-backed orioles were force-fed powder made from monarch butterflies collected in Santa Cruz, California, the jay ED$_{50}$ of which was previously estimated to be ~75 μg (ref. 14).

All birds were fed 130–140 mg of powder containing 146–157 μg of cardenolide, except one oriole that was fed half this dose. As controls, four additional orioles were force-fed 130–140 mg of ground Mexican butterfly abdomens, which contained low cardenolide content (35–38 μg). All 11 experimental orioles vomited, but none of the grosbeaks or control orioles became ill. These data show that the oriole is sensitive to the emetic effects of the cardenolide, but that the grosbeak is not.

The ED_{50} of the Mexican butterflies, combined with the data shown in Fig. 1a, shows that most of the Site Alpha monarch butterflies contain less than one emetic unit of cardenolide. The Mexico and California populations have a high proportion of butterflies with little or no cardenolide and contrast with the Massachusetts population in which no butterfly completely lacks poison. However, the cardenolides in both the Mexico and Massachusetts populations are much less emetic than those in the California butterflies. Consequently, the probability of any bird ingesting an emetic dose in Mexico is lower than in either of the other two populations (Fig 1b, c).

We offer three possible explanations for the high percentage of relatively non-emetic butterflies at Site Alpha. Total cardenolide decreases with time in overwintering monarch butterflies[21,22], and may result in increasing predation as the winter progresses. Second, Urquhart[23] has suggested that butterflies from different regions of the eastern breeding range may migrate to geographically distinct overwintering sites in Mexico. If this is correct, then other Mexican sites may contain more emetic butterflies reflecting regional differences in the local milkweed floras[24]. However, heavy bird predation has been observed at all known sites[1]. A third more general explanation results from preliminary TLC separations of the cardenolides in the Site Alpha butterflies. These suggest that a large proportion of the butterflies fed as larvae on *Asclepias syriaca* L. and/or *Asclepias speciosa* Torr., both of which contain low-potency cardenolides[9,25]. Agricultural and land use changes have increased the abundances of both plants in North America[24,26].

Our finding that individual monarch butterflies containing high concentrations of cardenolides are not protected against two major bird predators raises questions about the part played by these compounds in the butterflies' defence. It is possible that some of the overwintering butterflies which were rejected by the birds without being killed may be protected by other distasteful or toxic substances, such as pyrrolizidine alkaloids[27,28]. Alternatively, the chemical defence of the monarch butterfly may be ineffective in these overwintering sites and the only advantage to individuals within the colonies may derive solely from a reduced probability of being killed, comparable with the advantages of herding[1,29]. It is possible that the extensive predation is a recent phenomenon attributable to changes in the relative abundances of the milkweed species in eastern North America.

Table 2 Cardenolide concentrations (μg per 0.1 g dry weight of the wings) of monarch butterflies eaten to various degrees by black-headed grosbeaks at Site Alpha, January 1980

	Damage category and description	Sample size	Cardenolide concentration (μg per 0.1 g dry weight)		
			$\bar{x}$	s.d.	Range
a	Lacking abdomen and 2–3 wings	41	40	41	0–128
b	Lacking abdomen and 1 wing	19	42	42	0–129
c	Lacking abdomen only	30	42	47	0–207
d	Abdominal contents squeezed out, exoskeleton not eaten	9	20	26	0–77
e	Abdomen slightly squeezed	7	84	106	0–290
f	Thorax damage only	18	46	50	0–199
g	Lacking head only	10	78	65	0–186
h	Lacking 1–3 wings only	12	35	56	0–177
i	Moribund (unable to fly but with no visible damage)	31	33	51	0–169
	Entire sample	177	42	51	0–290

The butterflies are grouped into nine damage categories, the mean cardenolide concentrations of which do not differ significantly (one way ANOVA $F_{8,176} = 1.17$, $P = 0.32$). The absence of an obvious relationship between degree or kind of damage and cardenolide concentration indicates that once they have attacked monarchs, grosbeaks, unlike orioles, feed indiscriminately on them.

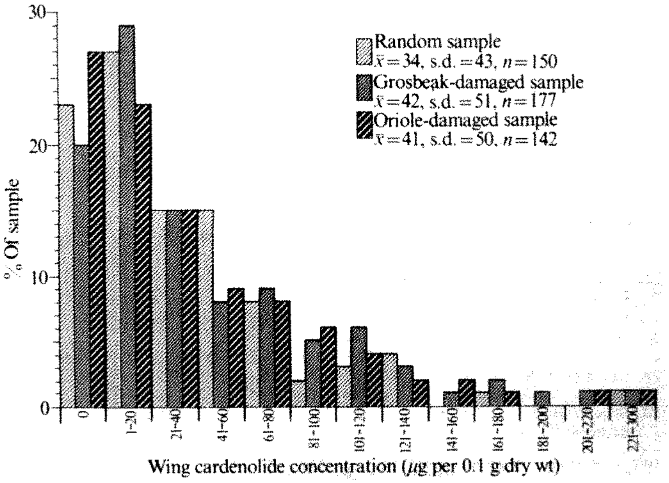


Fig. 2 Frequency distributions of cardenolide concentrations in three samples of overwintering monarch butterflies from Site Alpha in Mexico, January 1980. The oriole and grosbeak samples are partly eaten butterflies dropped to the ground by the birds and collected by us during their feeding bouts in the butterfly colony. The random sample consists of undamaged live butterflies collected within the colony. Two wings per butterfly were analysed. To verify that wing concentration reflects that of the body, we separately analysed the wings and intact bodies of 50 male and 50 female monarch butterflies and found a high correlation between μg of cardenolide per 0.1 g of wing and the total μg per wingless body; $n = 100$, $r = 0.88$, $F = 320$, $P < 0.0001$. The figure shows that the concentrations of the cardenolides in the butterflies damaged by both bird species are distributed similarly to those of the random collection. Because the three distributions are skewed, we subjected all data to $\log_{10}$ transformations and compared the three means by one-way analysis of variance ($F_{2,466} = 0.91$, $P = 0.40$). These data provide strong evidence that cardenolides give individual butterflies at Site Alpha no obvious protection from either orioles or grosbeaks.

This research was supported by NSF grant DEB78-10658 with L.P.B. as principal investigator, and by the Leslie T. Webster Fund of Amherst College. We thank E. Romero, J. and E. de Jesus, R. Barthelemy, L. Liscomb, B. Olds, S. Friedrich and R. Waide for help in Mexico, R. Lindley for liason, S. Glazier for laboratory analyses, P. Walford and B. Steele for computing assistance, J. Lane for the Santa Cruz sample, A. Krause and B. Tiffany for technical help, and K. Anderson and T. Lloyd-Evans of Manomet Bird Observatory and M. Clench and R. Leberman of Powdermill Nature Reserve for orioles and grosbeaks. We also thank K. Brown, W. Calvert, F. Chew, C. Nelson, T. Sargent, J. Seiber and N. Stamp for comments on the manuscript, and K. D. Sargent, M. Fenner and E. Belcher for artwork.

Received 10 November 1980; accepted 19 March 1981.

1. Calvert, W. H., Hedrick, L. E. & Brower, L. P. *Science* **204,** 847–851 (1979).
2. Reichstein, T., von Euw, J., Parsons, J. A. & Rothschild, M. *Science* **161,** 861–866 (1968).
3. Brower, L. P., Ryerson, W. N., Coppinger, L. L. & Glazier, S. C. *Science* **161,** 1349–1351 (1968).
4. Swynnerton, C. F. M. *Afr. orn. Un.* **11,** 32–108 (1915).
5. Parsons, J. A. *J. Physiol., Lond.* **178,** 290–304 (1965).
6. Dixon, C. A., Erickson, J. M., Kellett, D. N. & Rothschild, M. *J. zool Res.* **185,** 437–467 (1978).
7. Brower, L. P. in *Biochemical Coevolution* (ed. Chambers, K. L.) 69–82 (Oregon State University Press, Corvallis, 1970).
8. Rothschild, M. & Kellett, D. N. *J. Ent.* **46,** 103–110 (1972).
9. Brower, L. P., Fink, L. S. & Waide, R. B. (in preparation).

10. Chaney, S. G. & Kare, M. R. *J. Am. vet. med. Ass.* **149,** 938–943 (1966).
11. Benson, J. N. *et al. Toxicon* **17,** 155–165 (1979).
12. Brower, L. P., McEvoy, P. B., Williamson, K. L. & Flannery, M. A. *Science* **177,** 426–429 (1972).
13. Brower, L. P., Edmunds, M. & Moffitt, C. M. *J. Ent.* **49,** 183–196 (1975).
14. Brower, L. P. & Moffitt, C. M. *Nature* **249,** 280–283 (1974).
15. Brower, L. P. *Scient. Am.* **220,** 22–29 (1969).
16. Brower, L. P., Seiber, J. N., Nelson, C. N. & Bond, C. (in preparation).
17. Roeske, C. N., Seiber, J. N., Brower, L. P. & Moffitt, C. M. *Recent Adv. Phytochem.* **10,** 93–167 (1976).
18. West, D. A. *Auk* **79,** 399–424 (1966).
19. Brower, L. P. & Glazier, S. C. *Science* **188,** 19–25 (1975).
20. Brower, L. P., Brower, J. V. Z. & Corvino, J. M. *Proc. natn. Acad. Sci. U.S.A.* **57,** 893–898 (1967).
21. Tuskes, P. M. & Brower, L. P. *Ecol. Ent.* **3,** 141–153 (1978).
22. Brower, L. P., Calvert, W. H., Glazier, S. C. & Sheppard, M. (in preparation).
23. Urquhart, F. A. & Urquhart, N. R. *Can. J. Zool.* **56,** 1759–1764 (1978).
24. Woodson, R. E. Jr. *Ann. M. bot. Gdn* **41,** 1–211 (1954).
25. Brower, L. P., Seiber, J. N., Nelson, C. N., Lynch, S. P. & Tuskes, P. M. (in preparation).
26. Urquhart, F. A. & Urquhart, N. R. *Can. Fld Nat.* **93,** 41–47 (1979).
27. Rothschild, M. & Marsh, N. *Ent. exp. appl.* **24,** 437–450 (1978).
28. Edgar, J. A., Boppré, M. & Schneider, D. *Experientia* **35,** 1447–1448 (1979).
29. Edmunds, M. *Defence in Animals* (Longman, New York, 1974).

Eclosion hormone may control all ecdyses in insects

James W. Truman*, Paul H. Taghert,
Philip F. Copenhaver, Nathan J. Tublitz
& Lawrence M. Schwartz

Department of Zoology, University of Washington, Seattle,
Washington 98195, USA

The moulting process in insects is initiated and directed by the hormone 20-hydroxyecdysone[1], which causes the secretion of a new cuticle and the internal changes that are necessary for the new stage. When the new cuticle is complete, the old one is then shed or *ecdysed*. In Lepidoptera a neurosecretory peptide, eclosion hormone (EH)[2–4], triggers adult ecdysis by acting on the central nervous system to elicit the ecdysial motor programmes[5] and also causes other physiological changes associated with ecdysis[6]. Although it was originally thought that EH was used only for the rather specialized adult ecdysis[3], we report here that the hormone is used for all postembryonic ecdyses in the life history of moths, but its source varies with developmental stage. Moreover, EH may be common to a wide variety of insects. For the ecdyses of the adult[4], pupa[7] and the 5th instar larva of the tobacco hornworm, *Manduca sexta* (P.F.C., unpublished results), EH appeared in the blood before the start of the behaviour. Also, injections of partially purified EH caused precocious ecdysis of the three stages with 50% responses occurring at 0.3, 0.02 and 0.01 EH units, respectively.

Although EH seems to be associated with the ecdyses of the adult, pupa and 5th instar larva, the source of the blood-borne hormone varies. EH is distributed in the CNS between the brain and its neurohaemal organ, the corpora cardiaca (CC), and the ventral chain of segmental ganglia[8]. The roles of these two regions in the control of ecdysis were examined by determining their EH titres throughout the life history of *Manduca*.

As seen in Fig. 1, ~10^{-3} units of EH activity were extracted from heads or ventral ganglia at the time of hatching. Through the five larval instars, both tissues showed accumulation of hormone during the feeding period of each stage, followed by a plateau during the time of formation of the new cuticle for the next stage. In the case of the 5th instar, this plateau extended from the start of the wandering stage through the period of pupal cuticle production. At each ecdysis, a sharp drop was then seen in the hormone stores of the ventral ganglia. A drop was also observed in the hormone content of the head during ecdysis of the 2nd instar but no further depletions were detected in either whole heads or brains at subsequent larval ecdyses or at pupal ecdysis. Examination of the CC before pupal ecdysis revealed no detectable EH activity.

Adult development was marked by a gradual increase in the titre of hormone in the ventral ganglia. No depletion was seen at adult ecdysis and accumulation continued into adult life. Titre changes in the brain and CC during this time were more complex. Hormone levels increased in the brain through the first half of development but gradually declined during the second half. These changes were accompanied by a >100-fold increase in activity in the CC. This accumulation of hormone in the CC was apparently due to hormone transferred from the brain because when brains were removed from animals early in adult development and their CC later assayed for the presence of EH, the latter lacked any activity ($n = 5$). At adult eclosion the CC showed a 70% loss of activity. EH activity persisted in both structures through at least the first 4 days of adult life.

A major feature of the titre curves in Fig. 1 is the prominent depletion in stored hormone which occurs at each ecdysis. At the ecdyses of the adult, pupa and 5th instar larva, each drop coincides with the appearance of EH activity in the blood. Moreover, the magnitude of the depletions (0.7, 0.06 and 0.05 EH units, respectively) compares favourably with the dosages required to provoke ecdysis in the same stages. On the basis of these data we have concluded that the sudden drops observed at each of the six ecdyses represent a release of EH into the blood and therefore that all ecdyses in *Manduca* are triggered by EH.

A role for EH during embryonic development was determined using the relatively large embryos of the moth *Hyalophora cecropia*. Preliminary experiments indicated that EH was present in embryos, mainly in the head region. As seen in Fig. 2, significant quantities of EH activity appeared in the head on day 5 of development. The titre remained high for the next 2 days but then dropped sharply between days 7 and 8 and remained low until the time of hatching on day 10. This drop in stored hormone coincided with the ecdysis of an embryonic cuticle by the developing 1st instar larva within the egg. Thus, it seems that EH is involved in embryonic as well as postembryonic ecdyses.

Our data suggest that EH storage and release occurred in either of two regions of the nervous system—the brain–CC complex or the ventral ganglia. Partial chemical characterization of the ecdysis stimulating activity from both regions revealed a single peptide with properties extremely similar or identical to EH[8]. The two regions seem to function independently. Indeed, experiments on adult ecdysis of the giant silkmoths show that the brain controls EH release at this stage[2], whereas at an earlier stage (pupal ecdysis) in *Manduca*, the previous removal of the brain from wandering stage larvae did not interfere with the subsequent release of EH from the ventral ganglia[8]. These experiments, albeit on different moth species, together with the data in Fig. 1, suggest two anatomically distinct systems for the synthesis and release of EH. However, the final proof requires the identification of the EH neurones in both regions.

Release occurs from the ventral ganglia during all postembryonic ecdyses, except for that of the adult which is controlled by the brain–CC axis. In these ecdyses, the choice of release site is correlated with the nature of the control over EH release. The time of adult ecdysis is controlled by a circadian clock which, at least in the case of the silkmoths[2], is in the brain. This clock regulates adult ecdysis through the brain–CC system[3]. By contrast, there is no direct circadian control over the larval and pupal ecdyses, but they occur at a developmentally fixed time after the start of the moult[9]. These developmentally triggered ecdyses are controlled by the ventral ganglia neuroendocrine system. A circadian involvement in the embryonic ecdysis has not been determined. Thus it is not known whether the above correlation also holds in the embryo. The use of the brain system by the embryo may also be due to its being the earliest part of the CNS to mature[10]. The transition from brain regulation in the embryo to regulation by the ventral ganglia in larvae seems to occur gradually as ecdysis to the 2nd-stage larva uses both systems.

The search for EH in insects other than Lepidoptera has been suggestive but not conclusive[11,12]. However, in these studies

---

* To whom correspondence should be addressed.

0028-0836/81/190070—02$01.00

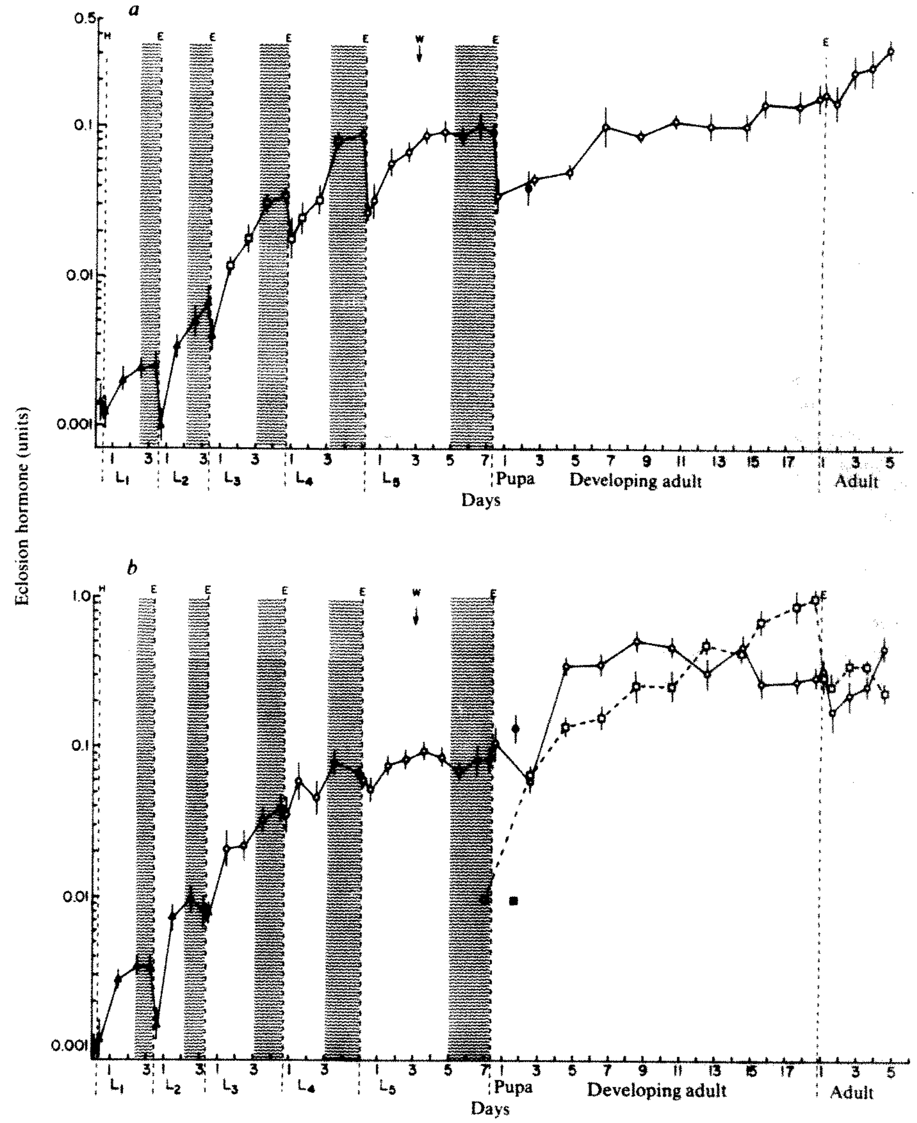


Fig. 1 Eclosion hormone activity extracted from the nervous system of *M. sexta* during its larval, pupal and adult stages. Depending on the stage, tissues from 5 to 300 animals were pooled, homogenized in saline with phenylthiourea, heat treated at 80 °C, centrifuged and the supernatant assayed for EH activity by injection into staged *Manduca* prepupae. The EH activity is expressed in units per animal (one unit being defined as the activity present in a CC pair from a *Manduca* on the last day of adult development). *a*, EH activity stored in the ventral ganglia. Symbols represent the hormone present in: △, mid-ventral strip of integument that included the segmental ganglia; □, abdominal and last two thoracic ganglia; ○, abdominal ganglia. *b*, EH activity present in cephalic structures. Symbols represent activity from extracts of: △, head and prothorax (the latter taken to insure inclusion of the CC); ○, brain and □, CC pair (the first two values represent the maximum amount of activity that could have been present but not detected in the conditions of the assays). At the times of hatching (H) and ecdyses (E), animals were selected ~4 h before (as judged by developmental markers) and 1 h after the event. W, start of wandering stage. The shaded areas represent the approximate time that new cuticle is being made in preparation for ecdysis (not indicated for adult development). The time marks along the abscissa correspond to lights-off for each day. Each determination is the mean (±s.e.) of at least nine assays. Solid symbols are from diapausing pupae.

only the brain was examined as a possible hormone source. The finding that the ventral ganglia controlled most of the ecdyses in *Manduca* prompted us to examine the ventral ganglia as well as the brain in a number of species. A preliminary survey revealed EH activity (as evidenced by the stimulation of precocious pupal ecdysis in *Manduca*[7]) in heat-treated extracts of nervous systems of juvenile forms from the hemimetabolous orders, Orthoptera (*Schistocerca*, *Melanopus*, *Periplaneta*) and Hemiptera (*Pyrrhocoris*), and the holometabolous orders, Coleoptera (*Dytiscus*), Trichoptera (*Dicosmoecus*) and Diptera (*Tipula*, *Calliphora*). These preliminary results should be regarded with caution because we have not yet shown that these activities from other orders are chemically similar to the moth EH, and that they are used to stimulate ecdysis in the insects in which they are found. Nevertheless, these results suggest that EH is present in a wide variety of insects and that it provides a general mechanism by which ecdyses are triggered in insects.

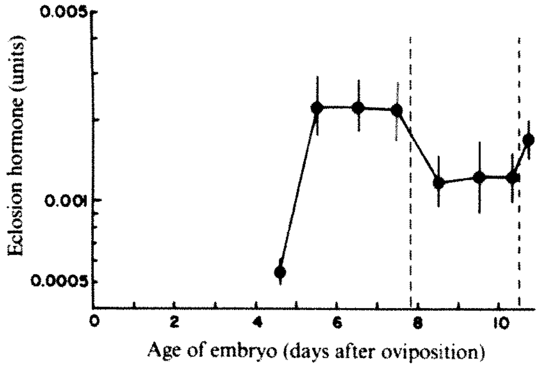


Fig. 2 Eclosion hormone activity extracted from the head and prothoracic segment of *Hyalophora cecropia* embryos during embryonic development. Activity is expressed as units per animal. Symbols represent the mean (±s.e.) of eight determinations. The time marks correspond to lights-off. The dashed vertical lines indicate the times of ecdysis of the embryonic cuticle and of hatching, respectively.

We thank Dr Shirley Reiss for technical assistance during part of this study and Professor Lynn M. Riddiford for helpful comments on the manuscript. The work was supported by NSF grant PCM 77-24878 and NIH grant R01 NS13079. P.H.T. and N.J.T. were supported by NSF predoctoral fellowships, L.M.S. by a NIH training program and J.W.T. by a NIH career development award.

Received 23 December 1980; accepted 3 March 1981.

1. Riddiford, L. M. & Truman, J. W. in *Insect Biochemistry* (ed. Rockstein, M.) 307–357 (Academic, New York, 1978).
2. Truman, J. W. & Riddiford, L. M. *Science* **167**, 1624–1626 (1970).
3. Truman, J. W. *Biol. Bull. mar. biol. Lab., Woods Hole* **144**, 200–211 (1973).
4. Reynolds, S. E., Taghert, P. H. & Truman, J. W. *J. exp. Biol.* **78**, 77–86 (1979).
5. Truman, J. W. *J. exp. Biol.* **74**, 151–174 (1978).
6. Truman, J. W. in *Insect Biology in the Future* (eds Locke, M. & Smith, D. S.) 385–401 (Academic, New York, 1980).
7. Truman, J. W., Taghert, P. H. & Reynolds, S. E. *J. exp. Biol.* **88**, 327–337 (1980).
8. Taghert, P. H., Truman, J. W. & Reynolds, S. E. *J. exp. Biol.* **88**, 339–349 (1980).
9. Truman, J. W. *J. exp. Biol.* **57**, 805–820 (1972).
10. Khan, T. R. & Fraser, A. in *Neurosecretion* (eds. Heller, H. & Clark, R. B.) 349–369 (Academic, London, 1962).
11. Carlson, J. R. *J. comp. Physiol.* **115**, 299–317 (1977).
12. Miller, P. L. & Mills, P. S. in *Perspectives in Experimental Biology* Vol. I. (ed. Spencer-Davis, P.) 199–208 (Pergamon, Oxford, 1976).

Positional signal transmission in the developing chick limb

Lawrence S. Honig*

Department of Biology as Applied to Medicine, Middlesex Hospital Medical School, London W1P 7PN, UK, and Division of Developmental Biology, National Institute for Medical Research, Mill Hill, London NW7 1AA, UK

* Present address: Laboratory of Developmental Biology, University of Southern California, GER-319, PO Box 77912, Los Angeles, California 90007, USA.

It has been suggested that the pattern along the antero-posterior axis of the embryonic developing chick limb arises from a gradient of diffusible morphogen[1,2] produced by a special region of cells on the posterior margin of the limb bud, called the polarizing region[3]. Grafts of posterior polarizing-region tissue to an anterior site in an embryonic chick wing bud result in mirror-image duplication of the limb bud; typically three extra wing digits are produced (Fig. 1*f*, *g*). Direct evidence for a long-range signal has, however, been lacking, and recently it has been suggested[4] that chick wing development might better be understood as resulting only from local, neighbour–neighbour-like interactions of the sort postulated to occur during regeneration of insect and amphibian limbs and insect imaginal disks[5]. As we report here, operations were performed in which chick wing buds were made host to grafts of leg bud responding tissue adjacent (posteriorly) to a grafted polarizing region (Fig. 2). The results showed that the pattern signal from the polarizing region can be transmitted through the leg tissue over distances of tens of cell diameters.

Double grafts in which leg tissue provides a living 'barrier' between grafted wing polarizing region and distal host chick wing tissue, are shown diagrammatically in Fig. 2. Leg tissue maintains its character in wings and when placed at wing bud tips develops into toes[6] which are recognizable through the patterns of their distinctively shaped cartilage elements and associated tissue (Fig. 1*h* and legend). (Many grafts also used quail marked tissue and the details of these will be presented elsewhere.) Sixty-six grafts were performed as described in Fig. 2 using leg anterior pieces ranging in width from 100 to 350 μm. In nearly all (61) cases the leg tissue, as expected, developed into recognizable toes. The important question was whether the signal from a grafted polarizing region could influence the wing tissue on the posterior side of the leg tissue barrier. In 34 cases (average width 215 ± 70 μm s.d.), only host wing digits were found posteriorly to toes (Fig. 1*e*), but in another 27 cases (average width 189 ± 57 μm s.d.) mirror-image duplicated wing digits appeared posterior to the toes (Fig. 1*a–d*). In general, leg anterior grafts of narrower width tended to yield limbs of this latter category, in which the posterior host tissue was influenced, while anterior grafts of large width provided more examples of posterior host tissue being unaffected.

Interpretations of these experiments depend on several considerations. First, control grafts of anterior leg tissue, without the additional polarizing region grafts, commonly did not give rise to any extra digits (Fig. 1*g*) in the resultant chick wings: out of 32 such operations, only 5 wings presented extra cartilage elements of which 4 possessed toe digits I and 1 had only an unidentifiable cartilage nodule. Therefore, it was the long-distance influence of the grafted polarizing region, not the interaction between the graft of the leg anterior tissue and the

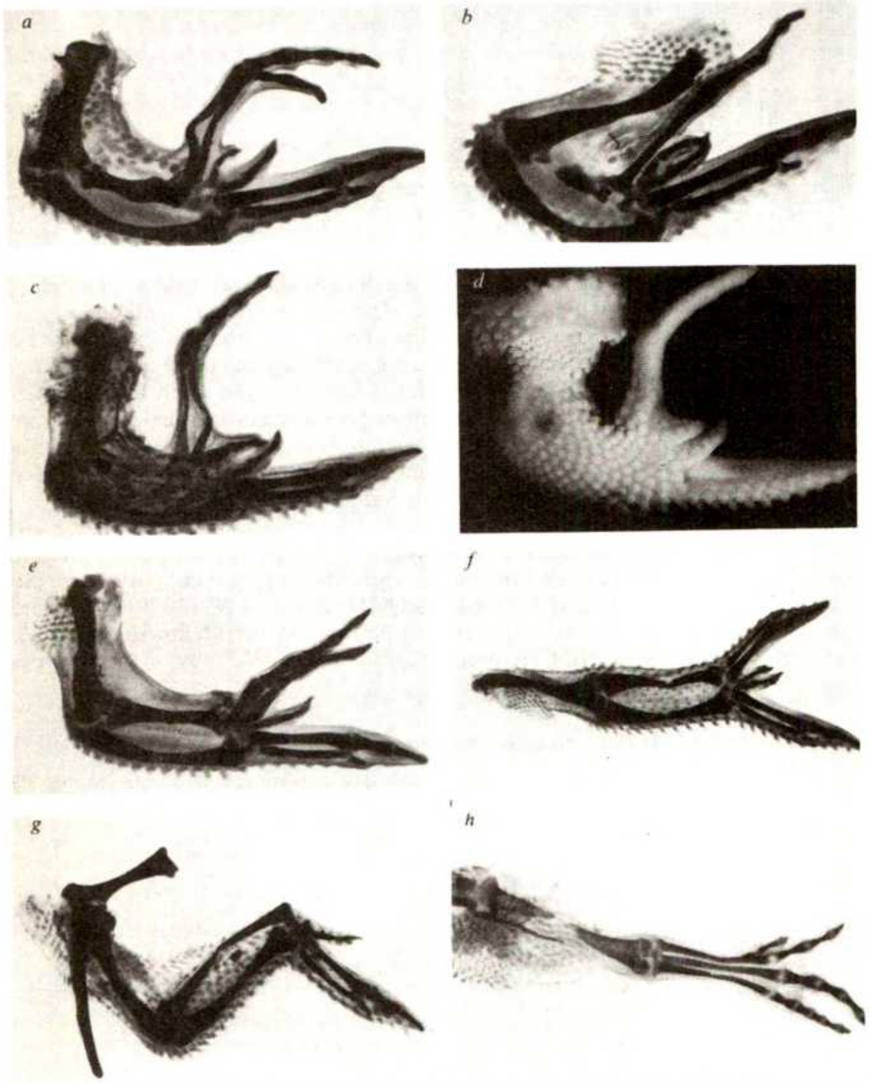


Fig. 1 Wholemount limbs from 10–12-day chick embryos (magnifications about ×6) following fixation, staining with alcian green 2GX, dehydration with ethanol and clearing with methyl salicylate[9]; limb *d* was not cleared. Wing digits are denoted by Arabic numerals; leg digits are represented by Roman numerals. Digit patterns are specified in anterior to posterior order. Toe digits (*h*) differ from wing digits (*g*) by the long thin metatarsals of II, III and IV and their more numerous phalanges. Toe I has some similarity to wing digit 2 but is distinguished by several features. Toe I normally arises not at the wrist but as an incomplete metatarsal of inverted triangular shape at distal metatarsal level. By contrast, metacarpal 2 arises at the wrist and is of thick rectangular shape with an anterior 'knob' at midlength. The first phalange of toe I is, in common with other proximal toe phalanges, of hour-glass shape: the basic cylindrical shape is constricted at the centre so as to resemble an 'X'; the first phalange of digit 2 is more like a narrow truncated cone whose proximal width is greater than that distally and is not constricted in the centre. The soft tissue of toe digits clings closely to the skeleton whereas it is more muscular and filled out for wing digits. At late stages, feather germs are present on avian wing digits but are not present on the toes which only develop distinctive scale primordia in these chicks and quail. *a–d*, Examples in which posterior host wing was influenced through a leg barrier: digit patterns are (*a*) III II 2 2 3 4, (*b*) II 2 2 3 4 and (*c*) 4 III 2 2 3 4. In *d*, specimen *c* is shown before it was cleared, showing the feather germs on the duplicate host digit (digit 2) and scales on the leg digit. *e*, Result of an operation in which posterior host wing was unaffected; the identifiable digit pattern is III II 2 3 4. *f*, Wing reduplication following graft of polarizing region only, with digit pattern 4 3 2 2 3 4. *g*, Wing with normal set of cartilage elements found after graft of leg anterior tissue alone, with digit pattern 2 3 4. *h*, Distal elements (tibia–fibula and foot) of normal control unoperated leg showing digit pattern I II III IV.

0028-0836/81/190072—02$01.00

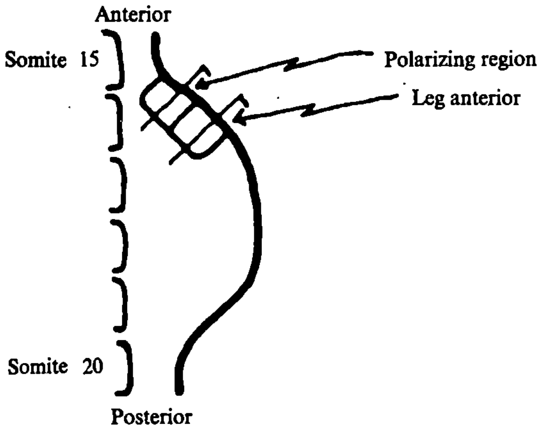


Fig. 2 Diagram of operation consisting of double graft, of wing polarizing region (donor stage 20–22) and leg anterior tissue (donor stage 19–23), each pinned with a platinum pin made from 25-μm diameter wire, replacing host wing bud tissue opposite somites 16/17. Hosts were chick embryos stage 19–21 and received several drops of Hank's balanced salts solution containing 20 mm HEPES and 1% antibiotic–antimycotic (Gibco) before being resealed with Sellotape and reincubated at 38 °C.

host, which caused the extra wing digits 2 obtained above, following the double grafts. Second, it is possible that polarizing region cells were physically migrating (in a posterior and distal direction) around the leg anterior tissue so as to contact directly the host tissue giving rise to the digits 2. This seems unlikely in the absence of any evidence for such migration in chick wing buds[7]. Nevertheless, control experiments were performed as in Fig. 2 but using anterior tissue from proximal parts of older (≥stage 22) limb buds. Such tissue has less competence to respond to or transmit the signal of the polarizing region and after such grafting operations 13 cases had no toes at all and 4 had a toe adjacent to host structures; in none of these operations were host structures duplicated. Third, although the host origin of duplicate wing digits 2 of the sort in Fig. 1*a–d* was indicated by their morphology and feather germs, it was confirmed by histological sections of limbs of this type resulting from operations using quail anterior tissue barriers. The duplicate digits 2 consisted almost entirely (>90%) of host chick cells.

The ability of polarizing region tissue to influence distal wing tissue a distance away from it is in direct distinction to expectations of local intercalation type models[4,5]. The experiments here show that the polarizing region can signal patterning information over a distance of about 200 μm, or about 20 cell diameters. But, there is no evidence for influencing positional information of cells at larger distances of over 300 μm. The results of several lines of investigation suggest that this distance is, at the stages operated on, roughly the antero-posterior extent of the presumptive limb field[8]: these include the absolute number of marked digits obtained using the double grafts described here (L.H., in preparation), fate maps[7], barrier experiments interpretable as fate maps[9], and grafts of two polarizing regions[10,11]. The long-range communication over ~20 cell diameters may be through a freely diffusible morphogen, but equally might occur through cell-to-cell transferred molecule(s), electric potentials, or even through a sequence of cell surface–surface interactions[12].

I thank Dr Lewis Wolpert for discussion and advice, and acknowledge the support of the Anna Fuller Fund (Fellowship No. 487) and the MRC.

Received 28 November 1980; accepted 20 February 1981.

1. Wolpert, L. *Curr. Topics dev. Biol.* **6,** 183–224 (1971).
2. Tickle, C., Summerbell, D. & Wolpert, L. *Nature* **254,** 199–202 (1975).
3. MacCabe, A.B, Gasseling, M.T. & Saunders, J.W. Jr *Mechanisms Ageing Dev.***2,** 1–12 (1973).
4. Iten, L.E. & Murphy, D.J. *Devl Biol.* **75,** 373–385 (1980).
5. French, V., Bryant, P.J. & Bryant, S.V. *Science* **193,** 969–981 (1976).
6. Saunders, J.W. Jr, Gasseling, M.T. & Cairns, J.M. *Devl Biol.* **1,** 281–301 (1959).
7. Stark, R.J. & Searls, R.L. *Devl Biol.* **33,** 138–153 (1973).
8. Summerbell, D. & Honig, L.S. *Am. Zool.* **21** (in the press).
9. Summerbell, D. *J. Embryol. exp. Morph.* **50,** 217–233 (1979).
10. Wolpert, L. & Hornbruch, A. *J. Embryol. exp. Morph.* (in the press).
11. Summerbell, D. *J. Embryol. exp. Morph.* (in the press).
12. Babloyantz, A. *J. theor. Biol.* **68,** 551–561 (1977).

Neurotensin blocks certain amphetamine-induced behaviours

Gregory N. Ervin, Larry S. Birkemo,
Charles B. Nemeroff & Arthur J. Prange Jr

Biological Sciences Research Center, Departments of Psychiatry,
Psychology and the Neurobiology Program,
University of North Carolina, School of Medicine, Chapel Hill,
North Carolina 27514, USA

Bilateral injections of either neurotensin (NT; 0.3, 1 or 5 μg in 1 μl artificial CSF) or haloperidol (HA; 2.5 or 5 μg in 1 μl 0.3% tartaric acid) into nucleus accumbens of rats markedly diminished the forward locomotion and rearing induced by d-amphetamine (AM; 2 mg per kg, IP). Neither NT nor HA affected the insomnia or sniffing component of AM arousal. Isovolumetric intra-accumbens injections of artificial CSF or the endogenous decapeptide, luteinizing hormone-releasing hormone (LHRH; 3 μg), did not affect AM behaviours. Since intra-accumbens injections of NT (1 μg) or HA (2.5 μg) neither altered forward locomotion or rearing observed in untreated rats placed in an open field nor a variety of reflex activities, the observed effects of NT and HA in AM-treated rats were probably not due to impaired motor function *per se*. In contrast, NT does not produce neuroleptic-like effects when injected into nucleus caudatus; HA (5 μg) blocked stereotyped sniffing, licking, biting and head bobbing observed after AM (5 mg per kg, IP), but NT (3 or 5μg) did not. Since NT and dopamine are present in substantial quantities in the nucleus accumbens, NT may act in the nucleus accumbens to modulate dopaminergic function.

Neurotensin (NT), a tridecapeptide (pGlu-Leu-Tyr-Glu-Asn-Lys-Pro-Arg-Arg-Pro-Tyr-Ile-Leu-OH), is an endogenous brain peptide which satisfies several neurotransmitter criteria[1,2] and may participate in both endocrine and behavioural functions[3,4]. Radioimmunoassay and immunohistofluorescence studies of rat brains indicate that certain diencephalic regions may have especially high concentrations of NT[5]. The nucleus accumbens is rich in both NT and dopamine (DA). Nucleus accumbens DA is contained in a rich plexus of nerve terminals that is an important part of the forebrain DA terminal system, which functions in arousal and exploration[6,7] and mediates increased forward locomotion and rearing stimulated by low doses of (+)amphetamine (AM)[8,9]. We describe here experiments which show that NT and the neuroleptic haloperidol (HA) have similar effects on AM-induced behaviour when injected into the nucleus accumbents of the rat, but that NT differs from HA in having no effect on AM behaviour when injected into the caudate. This suggests that NT may have a neuroleptic action without some of the side effects of drugs such as HA.

To test for possible interactions of NT and DA systems within the nucleus accumbens, we observed the effects of bilateral intra-accumbens microinjections of NT on the behavioural responses produced by a fixed dose of AM, and which are believed to be mainly dependent on DA release. Using a standardized technique of time sampling and behavioural scoring[9], we observed that NT attenuates AM-induced locomotion and rearing but not the associated insomnia or sniffing. Similar effects were also produced by bilateral intra-accumbens injections of HA, a DA receptor antagonist. We also studied the interactions of NT and the striatal DA system by comparing the effects of NT and HA on behavioural response to a high dose of

0028-0836/81/190073—04$01.00

Table 1 Effects of injections into nucleus accumbens on selected behavioural responses to amphetamine

	CSF	Neurotensin (μg)				LHRH (3 μg)	Tartaric acid	Haloperidol (μg)	
		0.1	0.3	1	5			2.5	5
Behavioural items (*n*)	30	10	9	11	15	8	10	10	10
Forward locomotion	15	8.5	6*	8*	5†	14.5	14	3†	1‡
	(1–36)	(2–28)	(0–28)	(0–20)	(0–21)	(1–30)	(2–33)	(0–8)	(0–30)
Rearing	4.5	2	1*	0*	0‡	4.5	1	1	0
	(0–38)	(0–26)	(0–14)	(0–11)	(0–11)	(0–13)	(0–4)	(0–3)	(0–11)
Sniffing	39.5	44	35	28	34	38.5	40	39.5	25*
	(10–53)	(18–50)	(27–40)	(11–46)	(12–54)	(19–50)	(30–51)	(26–48)	(0–49)
Sleeping	2.5	5	3	6	0	0	0	3	7.5
	(0–18)	(0–15)	(0–15)	(0–18)	(0–15)	(0–12)	(0–12)	(0–21)	(0–22)

Medians and ranges of behavioural scores over 3 h after (+)amphetamine (2 mg per kg i.p.) are shown. Immediately before amphetamine administration, all rats received injections into nucleus accumbens through indwelling cannulae. Doses shown were delivered bilaterally in 1 μl injectate over 1 min. Animals were observed for 30 s every 10 min for 3 h, and scored by a 'blind' observer using a standard sampling technique. The maximum possible score for each item was 54. Peptide scores were compared with artificial CSF scores, and HA scores with tartaric acid scores, by the Mann–Whitney *U*-test.

$*P<0.05$. $\dagger P<0.001$. $\ddagger P<0.01$.

AM acting on the striatum[8,10]. Although intracaudate injections of HA blocked AM-induced stereotyped behaviour, injections of NT did not.

Male Sprague–Dawley rats (350–450 g, Zivic Miller) were housed in individual cages at a temperature of ~23 °C with 12 h of light (0600–1800 h). Rats were cannulated under ketamine and phenobarbital anaesthesia when mounted in a stereotaxic instrument with the incisor bar 5 mm above the inter-aural line. Guide cannulae (22-gauge stainless steel), affixed to the cranium with dental acrylic, ended 3 mm dorsal to the sites of interest in the medial nucleus accumbens (9.6 mm anterior to inter-aural line, 1.25 mm right and left of sagittal sinus, and 6 mm ventral to dura) or caudate head (8.6 mm anterior to inter-aural line, 2.5 mm right and left, and 5 mm ventral to dura). These cannulae were implanted at least 2 weeks before behavioural testing. Stainless steel injection cannulae were 30 gauge. After testing, verification of accumbens cannulae placements was made in 70 out of 75 rats by frontally sectioning freshly frozen brains; 132 cannulae tracks terminated in the anteromedial nucleus accumbens, and eight were within 0.5 mm of this structure. Caudate cannulae placements were examined in 25 out of 30 rats; all cannulae tracks terminated within the nucleus caudatus.

On a test day, rats were placed individually in clear Plexiglass test cages (21 × 44 × 20 cm high) with wire mesh floors. Each test cage was mounted in a wooden box, with a smoked Plexiglass window (14 × 65 cm) which allowed viewing of a rat from ~20–40 cm. Test cages were lit with a 7.5 W lamp; the experimental room was dark. After an habituation period of 2 h, rats were removed from test cages and injection cannulae, connected by polyethylene tubing (PE 20) to a 10-μl Hamilton syringe and automatic dispenser, were inserted into nucleus accumbens or nucleus caudatus bilaterally; 1 μl of fluid was injected over 1 min into each site. Each injection contained 0.1, 0.3, 1.0 or 5.0 μg NT (in artificial cerebrospinal fluid (CSF), *p*H 7.4, Beckman); 3.0 μg luteinizing-hormone releasing hormone (LHRH, in CSF, Parke-Davis); 2.5 or 5.0 μg HA (in 0.3% tartaric acid, *p*H 3.0, MacNeil Laboratories); or CSF[11] or tartaric acid vehicles alone. Immediately after intra-accumbens injections, rats were injected intraperitoneally (i.p.) with 2 mg per kg AM (Smith, Kline and French), returned to test cages, and observed over the next 3 h by an observer ignorant of the treatments. Animals with cannulae in the nucleus caudatus were treated in the same way except that they received a higher dose of AM (5 mg per kg). Occurrence of the following six behavioural items was recorded in the rats with nucleus accumbens cannulae[6]: (1) forward locomotion—movements of the rat from one end of the cage to the other; (2) rearing; (3) sniffing—rapid and repetitive movement of the nostrils and/or snout; (4) sleeping—any resting posture with eyes closed throughout a whole 10-s interval; (5) rapid bobbing of the head; (6) grooming. In rats with nucleus caudatus cannulae receiving the higher dose of AM, the following behaviour types were considered to be stereotype indices: head bobbing (as described above), biting of the wire mesh floor, licking (of the floor or walls), and sniffing stereotype—continuous rapid and repetitive movement of nostrils and/or snout directed in patterns towards the floor or walls. Behaviour of each rat was observed for 30 s every 10 min for 3 h, and each 30-s observation was divided into three 10-s intervals. When a behaviour type was observed in a 10-s interval, that behaviour received a score of 1. Thus the total score for each behavioural item indicated the number of 10-s intervals in which that behaviour occurred, with a maximum possible score, after injection, of 54. It should be emphasized that these scores represent the incidence of behaviours; except for sniffing or sniffing stereotypes, we did not attempt to rate intensity of behaviours. Scoring of sniffing and sniffing stereotype was mutually exclusive. Using the Mann–Whitney *U*-test[12], we

Table 2 Effects of injections into caudate nucleus on selected behavioural responses to amphetamine

	CSF	Neurotensin (μg)		Tartaric acid	Haloperidol (5.0 μg)
		3.0	5.0		
Behavioural items (*n*)	13	10	10	12	10
Stereotypes (consisting of sniffing stereotype, head bobbing, biting and licking)	72	59.5	57	67	4.5*
	(46–98)	(32–92)	(14–98)	(45–99)	(0–39)
Forward locomotion	9	10	7	3	16*
	(0–53)	(1–33)	(0–16)	(0–26)	(1–52)
Rearing	2	3	3	1	7.5
	(0–14)	(0–32)	(0–9)	(0–26)	(0–37)
Sniffing	11	10.5	20.5	13	38.5*
	(2–32)	(2–25)	(0–47)	(4–19)	(28–47)
Sleeping	0	0	0	0	0
	(0)	(0)	(0)	(0)	(0–2)

Medians and ranges of behavioural scores over 3 h after (+)amphetamine (5 mg per kg i.p.) are shown. Immediately before amphetamine administration, all rats received injections into the caudate nucleus through indwelling cannulae. Doses shown were delivered bilaterally in 1 μl injectate over 1 min. Animals were observed for 30 s every 10 min for 3 h, and scored by a 'blind' observer using a standard sampling technique. Peptide scores were compared with artificial CSF scores, and HA with tartaric acid scores, by the Mann–Whitney *U*-test.

$*P<0.01$.

Table 3 Effects of injections into nucleus accumbens on open field performance

	Uncannulated		Cannulated		
	Uninjected	CSF	Neurotensin (1 μg)	Tartrate	Haloperidol (2.5 μg)
Behavioural items (*n*)	5	8	8	5	5
Squares traversed	91	101	101	107	112
	(62–215)	(51–249)	(14–208)	(42–367)	(21–188)
Rears	20	18	10	27	11
	(13–30)	(3–38)	(0–28)	(7–40)	(5–39)

Medians and ranges of behavioural scores of rats observed for 6 min in an open field are shown. On the basis of their performance in the open field on the previous day, cannulated rats were sorted to establish four comparable groups. Cannulated rats received injections into the nucleus accumbens 6 min before being put into the open field. Bilateral injections of NT (1 μg) and HA (2.5 μg) were given in 1 μl of artificial CSF and 0.3% tartaric acid, respectively. Statistical comparisons were made using the Mann–Whitney *U*-test. Rats injected with NT or HA did not differ from vehicle-injected or uncannulated rats.

compared scores from peptide-treated rats with scores from artificial CSF-treated rats, and HA scores with tartaric acid scores. (To determine if intra-accumbens vehicle injections alone affected AM-induced behaviours, we also observed nine cannulated rats which received AM but no intra-accumbens injections. Vehicle-treated animals were identical to non-injected animals except that tartaric acid significantly reduced amphetamine-induced rearing ($P < 0.05$).)

The four behavioural types affected by the low dose of AM in rats with nucleus accumbens cannulae are shown in Table 1. Nucleus accumbens injections of 0.3, 1.0 or 5.0 μg NT significantly attenuated the forward locomotion and rearing elicited by AM, but did not affect sniffing or insomnia. Injections of 2.5 or 5.0 μg HA also significantly decreased AM-induced locomotion, as described by others[13]. Injections of HA did not decrease insomnia, but the higher dose of HA (5.0 μg) significantly attenuated, but did not block, the sniffing component of AM arousal. The effects of NT on AM arousal were apparently not nonspecific attributes of neuropeptides because LHRH did not significantly alter AM-induced behaviours.

Table 2 shows the behavioural effects of the high dose of AM in rats with intracaudate cannulae. After AM administration, rats locomoted until stereotyped behaviours emerged and then prevailed. Caudate injections of HA (5.0 μg) blocked these stereotypes, and rats locomoted and sniffed as did rats given the lower dose of AM (Table 1). Such a shift in behavioural response to AM is also seen in rats with striatal DA denervations[8]. In contrast to the effects of intra-accumbens injections of NT, intracaudate injections of NT did not have the same effect as HA; intracaudate NT injections (3–5 μg) did not block stereotypes induced by a high dose of AM, but intrastriatal HA (5 μg) did, as described elsewhere[13].

We performed two evaluations which indicated that the effects of NT and HA on AM-induced locomotion and rearing were not due to generalized hypokinesia or motor impairment caused by intra-accumbens injections of NT or HA alone. Eight cannulated rats were examined for neurological functions daily for 4 days and showed stable behaviour. We observed hopping reaction in each limb, righting response on, or 30 cm above, a surface, vibrissae placing, chin placing, and the ability to right on an inverted metal grid[14]. Immediately after final baseline testing, rats were given intra-accumbens injections of 1 μg NT ($n = 3$), 2.5 μg HA ($n = 2$), or 1 μl CSF ($n = 3$). They were re-examined after 3.25 h by an observer ignorant of the treatments. None of the nucleus accumbens injections altered the performance of rats in these neurological tests (data not shown), suggesting no gross motor debilitation by NT or HA.

We observed open field (OF) performance of other intra-accumbens cannulated ($n = 26$) and uncannulated ($n = 5$) rats during 6-min sessions conducted on 2 consecutive days. A clear Plexiglass OF (80 × 80 × 60 cm high) was placed in a wooden box having a smoked Plexiglass top which allowed viewing of rats from ~90 cm. A grey piece of paper was placed under the OF; black lines drawn on the paper divided the floor into 64 squares 10 × 10 cm. The box was illuminated by two 7.5 W lamps; the experimental room was dark. Rats were placed in the OF facing the corner, and the following two behavioural items recorded: squares traversed—number of squares through which the rat moved all four limbs, and rears—vertical extension of forelimbs in any part of the OF. On the basis of OF performance on day 1, cannulated rats were divided into four equivalent subgroups which did not differ from uncannulated rats (data not shown). On day 2, cannulated rats received intra-accumbens injections containing 1 μg NT ($n = 8$), 2.5 μg HA ($n = 5$), 1 μl CSF ($n = 8$) or 1 μl tartaric acid ($n = 5$) 6 min before being placed in the OF. Scores of NT- or HA-treated rats did not differ significantly from scores of vehicle-treated or uncannulated rats (Table 3), which suggests that attenuation of AM-induced locomotion and rearing by intra-accumbens NT and HA was not due to an inability of rats to execute locomotor or rearing movements.

The neuroleptic-like effect of intra-accumbens NT on AM-induced behaviours is not the only evidence for possible NT–DA interactions in the brain. Like neuroleptics, NT delivered into the ventricular system elicits hypothermia[15,16], muscle relaxation[17], and potentiation of pentobarbital-induced narcosis[15]. Of particular interest is the recent report[18] that treatment of rats with HA produces marked increases in immunoreactive NT levels in the nucleus accumbens, with only a small increase in the striatum. Furthermore, intracerebrally administered NT has recently been shown to elicit an effect in two commonly used screens for neuroleptic agents, electrical self-stimulation and active avoidance responding[19]. Taken together these results suggest that NT may possess anti-dopaminergic (neuroleptic-like) properties[20]. However, because the precise central synaptic sites where NT and neuroleptics produce these various effects are unknown, the principle of central NT–DA interactions is still tentative. In view of the known anatomy of NT and DA systems, this work suggests that exogenous NT produces its effects at sites where endogenous NT is plentiful (for example, nucleus accumbens), and not where it is scarce (for example, nucleus caudatus) although DA may be concentrated at both sites. Further support for an NT–DA interaction in the nucleus accumbens has recently been provided by the finding that NT, applied iontophoretically, exerts a potent inhibitory effect on DA-sensitive neurones in the nucleus accumbens[21]. Because most clinically useful antipsychotic agents act at all DA receptor sites, they produce various neurological side effects presumably mediated by the nigrostriatal DA system. The finding that NT blocks AM-induced behaviours at a mesolimbic DA site but not at the extra-pyramidal DA sites, suggests the possibility of developing new pharmacological agents with antipsychotic efficacy, but without the unwanted side effects.

A preliminary account of certain of these data has appeared elsewhere[20]. This work was supported by grants from USPHS.

Received 17 November 1980; accepted 28 January 1981.

1. Uhl, G. R. & Snyder, S. H. *Life Sci.* **19**, 1827–1832 (1976).
2. Uhl, G. R., Goodman, R. R. & Snyder, S. H. *Brain Res.* **167**, 77–91 (1979).
3. Carraway, R. & Leeman, S. E. *J. biol. Chem.* **248**, 6851–6854 (1973).
4. Bissette, G., Manberg, P., Nemeroff, C. B. & Prange, A. J. Jr *Life Sci.* **23**, 2173–2182 (1978).

5. Kobayashi, R., Brown, M. & Vale, W. *Brain Res.* **126,** 548–564 (1977).
6. Young, R. C., Ervin, G. N. & Smith, G. P. *Pharmac. Biochem. Behav.* **5,** 565 (1976).
7. Fink, J. S. & Smith, G. P. *J. comp. physiol. Psychol.* **93,** 34–65, 66–84 (1979).
8. Kelly, P. H., Seviour, P. W. & Iversen, S. D. *Brain Res.* **94,** 507–522 (1975).
9. Ervin, G. N., Fink, J. S., Young, R. C. & Smith, G. P. *Brain Res.* **132,** 507–520 (1977).
10. Kennedy, L. A. & Zigmond, M. J. *Brain Res.* **168,** 408–413 (1979).
11. Merlis, J. K. *Am. J. Physiol.* **131,** 67–81 (1940).
12. Goldstein, A. *Biostatistics: an Introductory Text* (Macmillan, New York, 1964).
13. Pijnenburg, A. J. J., Honig, W. M. M. & Van Rossum, J. M. *Psychopharmacologia* **41,** 87–95 (1973); **45,** 65–79 (1975).
14. Sechzer, J. A., Ervin, G. N. & Smith, G. P. *Expl Neurol.* **41,** 723–737 (1973).
15. Nemeroff, C. B. *et al. Brain Res.* **128,** 485–496 (1977).
16. Bissette, G., Nemeroff, C. B., Loosen, P. T., Prange, A. J. Jr & Lipton, M. A. *Nature* **262,** 607–609 (1976).
17. Osbahr, A. J. III, Nemeroff, C. B., Manberg, P. J. & Prange, A. J. Jr *Eur. J. Pharmac.* **54,** 299–302 (1979).
18. Govoni, S., Hong, J. S., Yang, H-Y. T. & Costa, E. *J. Pharmac. exp. Ther.* **215,** 413–417 (1980).
19. Luttinger, D., Wiggins, R., King, R. A., Nemeroff, C. B. & Prange, A. J. Jr *Neurosci. Abstr.* **6,** 843 (1980).
20. Nemeroff, C. B. *Biol. Psychiat.* **15,** 283–302 (1980).
21. McCarthy, P. S., Walker, R. J., Yajima, H., Kitagawa, K. & Woodruff, G. N. *Gen. Pharmac.* **10,** 331–333 (1979).

Selective effects of somatostatin-14, -25 and -28 on *in vitro* insulin and glucagon secretion

Lawrence Mandarino, Debra Stenner, Wayne Blanchard, Steven Nissen & John Gerich*

Endocrine Research Unit, Departments of Medicine and Physiology,
Mayo Medical School and Mayo Clinic, Rochester,
Minnesota 55901, USA

Nicholas Ling, Paul Brazeau, Peter Bohlen, Frederick Esch & Roger Guillemin

Laboratories for Neuroendocrinology, Salk Institute for Biological
Studies, San Diego, California 92138, USA

The widely occurring[1–3] tetradecapeptide somatostatin (SRIF-14) has been variously implicated[4–8] as a neurotransmitter, a neurohormone, a cybernin (local regulatory factor) and a hormone. In the first isolation of SRIF-14 from hypothalamic extracts[9,10] and subsequent extracts of other tissues[11–15], peptides of higher molecular weight but with similar activity have been noted. Recently two such peptides have been characterized as the 28-amino acid SRIF-28 (from porcine gastrointestinal tract[16] and porcine and ovine hypothalamus[17,18]) and the 25-amino acid SRIF-25 (from ovine hypothalamus[17]), each of which consists of an N-terminal extension of SRIF-14. We now report that SRIF-28 and SRIF-25 are more potent than SRIF-14 in the inhibition of insulin release, but that SRIF-14 preferentially inhibits glucagon release. This suggests that SRIF-28 and SRIF-25 are not mere biosynthetic precursors of SRIF-14 and that their differential release may be physiologically important.

To compare the relative effects of SRIF-28, SRIF-25 and SRIF-14 on secretion of insulin and glucagon *in vitro*, each peptide was synthesized by solid phase methods[19]; their primary structure was recharacterized on a Beckman 8900 amino acid sequencer, and purity confirmed by HPLC. The peptides were dissolved in pancreas perfusate[20] and infused by side-arm syringe (0.1 ml min^{-1}) starting 10 min before and continuing for 5 min after a 15 min pancreatic stimulation by arginine (10 mM). Each pancreas was exposed to only a single concentration of a peptide.

Pancreases were isolated from overnight fasted Wistar rats (220-280 g) and perfused *in vitro* with a non-recirculating medium (3 ml min^{-1}) as described elsewhere[20]. Portal venous effluent was collected at 1-min intervals into chilled tubes containing 25 μM EDTA; after completion of experiments, the effluent was aliquoted and frozen until determination of insulin[21] and glucagon[22] contents by radioimmunoassay. Hormone secretion, presented as the total amount of hormone released during each 15-min stimulation period (see Table 1), was calculated by summation of the amount of the hormone released for each minute of the 15-min stimulation. Basal glucagon (~ 0.1 ng min^{-1}) and insulin (~ 1.5 ng min^{-1}) secretory rates during individual experiments were not significantly different.

SRIF-25 and SRIF-28 did not differ in their biological effects. At all concentrations studied, both peptides inhibited secretion of insulin more potently than did SRIF-14. Linear regression analysis indicated that 50% inhibition of insulin occurred at $\sim 1.4 \times 10^{-11}$ M for SRIF-28 and SRIF-25 and at $\sim 1.5 \times 10^{-10}$ M for SRIF-14. In contrast, SRIF-14 inhibited secretion of glucagon more potently than did SRIF-25 and SRIF-28; 50% inhibition of glucagon release occurred at $\sim 3.3 \times 10^{-12}$ M SRIF-14 and at $\sim 3.2 \times 10^{-11}$ M SRIF-25 and SRIF-28. Thus, both SRIF-25 and SRIF-28 seem to inhibit preferentially insulin release whereas SRIF-14 seems to inhibit preferentially glucagon secretion. As a non-recirculating perfusion system was used in these experiments, the different activities observed for these peptides cannot be explained by their differing rates of degradation.

It has been suggested that SRIF-28 may represent prosomatostatin, the biosynthetic precursor of SRIF-14[16,18,23]. However, prohormones have been found to be considerably less potent than their product[24]. Thus, the present demonstration that SRIF-25 and SRIF-28 are about 10-fold more potent than SRIF-14 in inhibiting insulin *in vitro* and previous reports[17,28] that synthetic replicates of these peptides are 10-fold more potent in inhibiting growth hormone secretion *in vitro* and *in vivo* suggest that SRIF-25 and SRIF-28 are not mere biosynthetic precursors for SRIF-14 in the classical sense of a prohormone but rather that they should be considered as alternate forms of biologically active somatostatins.

As even the secreted peptides may be further processed before interaction at receptors in target tissues[25], it will remain uncertain which peptide is the precursor and which the ultimate product until it is determined which peptide is actually bound at the SRIF receptor *in vivo*. Furthermore, it is possible that, as with proopiocortin which is processed differently in adenophyseal cells and cells of the pars intermedia[26], prosomatostatin may also be processed differently in various tissues. In some tissues, SRIF-14 may represent the predominant secretory product whereas in others it may be SRIF-28 or SRIF-25. Moreover, the same cell may release different amounts of the peptides in different physiological or pathological conditions.

There is evidence that both SRIF-14 and SRIF-28 are present in the pancreas, presumably in islet D-cells[27]. That SRIF-14 is a

Table 1 Effects of somatostatin-14, -28 and -25 on *in vitro* insulin and glucagon secretion

	Insulin (ng)	Glucagon (ng)
Control (12)	483 ± 13	57.7 ± 4.8
Somatostatin-14		
3.3×10^{-13} (14)	489 ± 28	39.9 ± 2.7*
3.3×10^{-12} (12)	389 ± 26*	27.2 ± 1.3*
3.3×10^{-11} (7)	308 ± 24*	19.8 ± 1.8*
1.7×10^{-10} (11)	227 ± 14*	12.2 ± 0.7*
Somatostatin-28		
3.3×10^{-14} (5)	449 ± 38	55.1 ± 5.7
3.3×10^{-12} (8)	276 ± 20*†	40.0 ± 3.6*†
3.3×10^{-11} (6)	210 ± 18*†	35.0 ± 1.1*†
1.7×10^{-10} (6)	157 ± 26*†	19.2 ± 2.0*†
Somatostatin-25		
3.3×10^{-14} (5)	452 ± 31	53.6 ± 1.5
3.3×10^{-12} (6)	273 ± 22*†	40.8 ± 4.7*†
3.3×10^{-11} (5)	239 ± 20*†	29.8 ± 2.1*†
1.7×10^{-10} (5)	166 ± 20*†	19.4 ± 2.9*†

Data represent the total amount of hormone released during each 15-min stimulation period (see text for details) and are given as mean ±s.e.m. Numbers in parentheses indicate the number of experiments.

* $P < 0.05$ against control.

† $P < 0.05$ against somatostatin-14.

---

* To whom reprint requests should be addressed.

more potent inhibitor of arginine-stimulated glucagon release than of insulin release has been previously noted[20]. The present finding that SRIF-28 is a more potent inhibitor of arginine-stimulated insulin release than of glucagon release raises the possibility that preferential secretion of SRIF-14 and SRIF-28 may provide a novel mechanism for selective regulation of insulin and glucagon release. This suggestion is consistent with the observation that SRIF-28 seems to comprise less than 5% of SRIF-like immunoreactivity within the pancreas[27] because SRIF-28 is a 10-fold more potent inhibitor of insulin release than is SRIF-14, the predominant molecular species in the pancreas. It remains to be determined, however, whether the present observations regarding arginine-stimulated insulin and glucagon release are applicable to other secretagogues of these hormones.

This work was supported in part by grants AM20837, AM07147, AM07352 and AM11811 from the NIH and from the Mayo Foundation, Hearst Foundation and the National Foundation March of Dimes.

Received 2 February; accepted 9 March 1981.

1. Hokfelt, T. *et al. Acta endocr.* **80,** 1–41 (1978).
2. Parson, J., Erlandsen, S., Hegre, O., McEvoy, R. & Elde, R. *J. Histochem. Cytochem.* **24,** 872–882 (1976).
3. Yamada, Y., Ito, S., Matsubara, Y. & Kobayashi, S. *Tohoku J. exp. Med.* **122,** 87–92 (1977).
4. Renald, L., Martin, J. & Brazeau, P. *Nature* **255,** 233–235 (1975).
5. Cohn, M. & Cohn, M. *Brain Res.* **96,** 138–141 (1975).
6. Guillemin, R. & Gerich, J. *A. Rev. Med.* **27,** 379–388 (1976).
7. Creutzfeldt, W. & Arnold, R. *Metabolism* **27,** 1309–1316 (1978).
8. Raptis, S., Schlegel, W., Lehmann, E., Dollinger, H. & Zoupas, C. *Metabolism* **27,** 1321–1328 (1978).
9. Brazeau, P. *et al. Science* **179,** 77–79 (1973).
10. Schally, A. *et al. Biochemistry* **15,** 509–514 (1976).
11. Patel, Y. & Reichlin, S. *Endocrinology* **102,** 523–530 (1978).
12. Larsson, L. *et al. Lancet* **i,** 666–668 (1977).
13. Vale, W. *et al. Metabolism* **25,** 1491–1494 (1976).
14. McIntosh, C. *et al. Gut* **19,** 656–663 (1978).
15. Arimura, A., Sato, H., DuPont, A., Nishi, N. & Schally, A. *Science* **189,** 1007–1008 (1975).
16. Pradayrol, L., Jornvall, H., Mutt, V. & Ribet, A. *FEBS Lett.* **109,** 55–58 (1980).
17. Brazeau, P. *et al. C.r. hebd. Séanc Acad. Sci., Paris* **290,** 1369–1372 (1980).
18. Schally, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **77,** 4489–4493 (1980).
19. Ling, N. *et al. Biochem. biophys. Res. Commun.* **95,** 945–951 (1980).
20. Gerich, J., Lovinger, R. & Grodsky, G. *Endocrinology* **96,** 749–754 (1975).
21. Herbert, V., Lau, K., Gottlieb, C. & Bleicher, S. *J. clin. Endocr. Metab.* **25,** 1375–1384 (1965).
22. Faloona, G. & Unger, R. in *Methods of Hormone Radioimmunoassay* (eds Jaffe, G. & Behrman, H.) (Academic, New York, 1974).
23. Meyers, C., Murphy, W., Redding, T., Coy, D. & Schally A. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
24. Steiner, D. in *Peptide Hormones* (ed. Parsons, J.) (Macmillan, London, 1976).
25. Peart, W. *Pharmac. Rev.* **17,** 143–182 (1965).
26. Eipper, B. & Mains, R. *J. Supramolec. Struct.* **8,** 247–262 (1978).
27. Benoit, R., Bohlen, P., Ling, N. & Guillemin, R. *Endocrinology* **107,** 2127–2129 (1980).
28. Brazeau *et al. Reg. Peptides* **1,** 255–264 (1981).

γ Interferon induced by *S. aureus* protein A augments natural killing and ADCC

William J. Catalona, Timothy L. Ratliff
& Robert E. McCool

The Yalem Research Laboratories of The Jewish Hospital of
St Louis and the Division of Urology, Washington University School
of Medicine, St Louis, Missouri 63110, U.S.A.

***Staphylococcus aureus* produces a protein identified as protein A (SpA) with a molecular weight of 41,000 (ref. 1) which binds to the Fc region of many types of mammalian IgG[2], activates complement[3], blocks opsonins[4] and is both chemotactic[5] and mitogenic[6,7]. SpA has been used for various immunological purposes including removal of immune complexes[8], arming effector cells with antibody[9], distinguishing between antibody-dependent cell-mediated cytotoxicity (ADCC) and natural killing (NK)[10], and exerting anti-tumour activity[11]. We have now demonstrated that SpA induces the production of γ interferon (IFN-γ) in nonadherent, T lymphocyte-depleted (E$^-$), Fc receptor-bearing (FcR$^+$) human peripheral blood lymphocytes. Our data also suggest that IFN-γ is a potent stimulator of both NK and ADCC activity.**

Table 1 Antigenic and physicochemical characterization of SpA-induced IFN

	Treatment of supernatant	Interferon (unit ml^{-1})
Physicochemical characterization	None (mock interferon)	≤5
	None (interferon)	250
	Trypsin	≤5
	*p*H 2	≤5
	56 °C	≤5
Antigenic characterization	None	2,000
	Anti-hu IFN-α	2,000
	Anti-hu IFN-γ	100

Human PBLs were treated with SpA (50 μg ml^{-1}) for 2 h, washed and incubated for an additional 24 h. Supernatants were collected as interferon preparations. Mock interferon was prepared in an identical fashion except that PBLs were treated with SpA diluent only. Supernatant treatments: trypsin: supernatants were treated with 100 μg ml^{-1} trypsin for 30 min at 37 °C; *p*H 2: supernatants were dialysed against glycine-HCl buffer (*p*H 2.0) for 24 h. The *p*H was returned to 7.2 by dialysis against minimal essential medium with Earl's salts; 56 °C: Supernatants were maintained at 56 °C for 1 h. Neutralization studies with antiserum against human (hu) IFN-α and hu IFN-γ were kindly performed by Dr John Stanton, University of Texas Medical Branch, Galveston, Texas.

Previous reports have shown that staphylococcal enterotoxin A (SEA), a T-cell mitogen, induces IFN-γ production[12]. However, the interferon-producing potential of SpA has not previously been reported, and has not hitherto been a consideration in the interpretation of the biological properties of SpA.

In the present study, human peripheral blood lymphocytes (PBLs) were isolated by centrifugation of heparinized whole blood on Ficoll–Hypaque gradients as described previously[13]. Interface cells were washed three times with phosphate-buffered saline (PBS) and resuspended to the desired concentration in RPMI 1640 medium supplemented with 10% fetal calf serum and 1% antibiotic–antimycotic solution (Gibco) (complete medium). Interferon was induced by incubating PBLs (5×10^6 ml^{-1}) with 50 μg ml^{-1} SpA (Pharmacia; no SEA contamination has been detected in SpA preparations from Pharmacia [James L. Richey (Pharmacia), personal communication]) for 2 h at 37 °C in a 5% CO_2 atmosphere. Following incubation with SpA, PBLs were washed three times with PBS, resuspended to 5×10^6 cells ml^{-1} in complete medium and incubated for 24 h at 37 °C. Supernatants were collected as interferon preparations and frozen at −70 °C until used. Mock interferon was prepared in an identical fashion except that SpA was not added to PBL cultures. Supernatants were assayed for interferon activity by measuring their capacity to inhibit the cytopathic effect of vesicular stomatitis virus on human foreskin fibroblast cultures (Biofluids). Titres are reported as International Reference units (NIH Human Leukocyte Reference Standard GO 23-902-527).

Interferon assays were performed on supernatants of SpA-treated PBLs from more than 20 donors. The interferon titres varied between donors, ranging from 100 to 300 U ml^{-1} (mean = 160 ± 20 U ml^{-1}) using unfractionated PBLs. A representative experiment is shown in Table 1. The results show that supernatants from SpA-treated PBLs exhibited antiviral activity whereas none was present in mock interferon preparations. Physicochemical and antibody neutralization studies (Table 1) suggest that SpA-induced interferon is immune interferon (IFN-γ). Cell fractionation studies (Table 2) showed that depletion of adherent cells on glass surfaces and removal of the majority of T lymphocytes by E-rosette depletion enhanced interferon production, whereas depletion of FcR$^+$ cells on antibody-sensitized sheep erythrocytes abrogated interferon production. These results suggest that with SpA, unlike other mitogen interferon inducers, macrophage depletion enhances

0028–0836/81/190077—03$01.00

Table 2 Determination of the mononuclear cell subpopulation responsible for SpA-induced interferon production

Experiment no.	Cell fraction	Interferon (units ml^{-1})	% Latex positive	Cell markers % FcR positive	% E-rosette positive
1	Unfractionated	250	18	ND	ND
	Macrophage enriched	125	64	ND	ND
	Macrophage depleted	1,580	2	ND	ND
2	Macrophage depleted	625	ND	5	ND
	FcR negative	25	ND	1	ND
3	Macrophage depleted	200	ND	ND	70
	T-lymphocyte depleted	1,250	ND	ND	16

The measurement of latex ingestion, Fc receptors and E-rosette formation was performed as previously described[13]. PBLs were incubated on glass Petri plates for 1 h at 37 °C. Macrophage-depleted cells were removed by vigorous washing. Macrophage-enriched cells were removed from the glass with a rubber policeman. For FcR-negative cells, FcR-positive lymphocytes were removed by IgG-coated sheep erythrocytes as previously described[13]. For T lymphocyte-depleted fraction, PBLs were incubated with sheep erythrocytes overnight at 4 °C. After incubation T lymphocytes were removed by Ficoll–Hypaque density centrifugation as previously described[13]. ND, no data.

rather than reduces interferon production[14] and further, that nonadherent, non-T, FcR^+ lymphocytes produce interferon after SpA stimulation.

Interferons have been classified as IFN-α, IFN-β and IFN-γ, based on antigenic and physicochemical properties. IFN-α, -β and -γ are antigenically distinct compounds[14]. Furthermore, IFN-α and IFN-β, unlike IFN-γ, are stable to pH 2 and to heating at 56 °C for 1 h (ref. 15). All interferon types are inactivated by trysin treatment. Our characterization studies revealed that SpA-induced interferon was inactivated by pH 2 treatment for 24 h or heating at 56 °C for 1 h. Moreover, treatment of SpA-induced interferon with antiserum to human IFN-α failed to neutralize its antiviral activity whereas antiserum to human IFN-γ succeeded, suggesting that SpA-induced interferon is IFN-γ.

Previous reports have shown that IFN-α and IFN-α inducers augment both ADCC and NK activity[16]. We therefore evaluated the effect of SpA-induced IFN-γ-containing supernatants on ADCC and NK activity (Fig. 1). Effector PBLs were treated with SpA-induced interferon-containing supernatants overnight at 37 °C (controls were treated with mock interferon) before the addition of ^{51}Cr-labelled target cells. The cytotoxic activity of both NK directed against K562 human myeloid leukaemia target cells and ADCC directed against Chang human liver cell targets was augmented by interferon-containing supernatants (Fig. 1). Significant augmentation was observed at effector/target cell ratios ranging from 6 : 1 to 50 : 1 (data not shown). Furthermore, trypsin digestion, heating at 56 °C and pH 2 treatment which inactivated the antiviral activity of SpA-induced interferon also significantly diminished that augmentation of cytotoxic activity (data not shown).

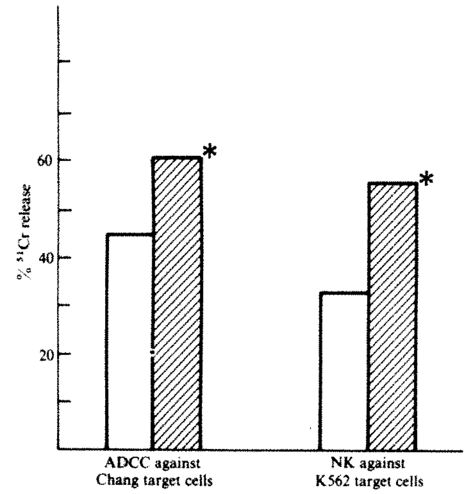


Fig. 1 Effect of SpA-induced interferon-containing supernatant on NK and ADCC activity of human PBLs. Both ADCC and NK assays were performed at 50 : 1 effector/target ratios as described previously[13]. Open bars represent untreated effector cells, hatched bars interferon-treated effector cells. *, $P < 0.005$ by Student's t-test.

SpA has often been used in short-term cytotoxicity assays to distinguish between ADCC activity, which is inhibited by SpA, and NK activity which is not[17]. Kinetic studies on the rate of interferon production after SpA treatment showed that interferon activity was not apparent until 8 h after SpA treatment and peaked by 24 h (data not shown). Thus, only long-term cytotoxicity assays would be expected to reflect the interferon-induced augmentation of NK and ADCC activity. Indeed, Pape *et al.*[18] recently reported that SpA-treated PBLs exhibited augmented NK activity against T 24 bladder cancer cells in an 18-h assay, but not against K562 cells in a 4-h assay. Other investigators have observed that SpA-treated effector cells exhibited augmented NK activity in 18-h, but not in short-term[10], assays[17].

Previous investigators have not considered the interferon-inducing potential of SpA in interpreting its diverse effects in biological systems. Cowan *et al.*[8] added SpA to ADCC assays containing immune complexes and attributed the enhanced cytotoxicity observed to removal of the immune complexes. Sulica *et al.*[9] found that effector cells treated with target cell-specific antibody complexes to SpA exhibited greater cytotoxic activity than effector cells treated with antibody alone. They attributed the enhanced cytotoxicity to increased antibody arming in the presence of SpA. Our results suggest that the enhancement of cytotoxicity observed by both Cowan *et al.*[8] and Sulica *et al.*[9] may be attributed, at least in part, to interferon induction.

Ray and co-workers[11] demonstrated that *S. aureus* (Cowan), a producer of SpA, has antitumour activity against a murine methylocholanthrene-induced fibrosarcoma. These investigators suggested that the antitumour activity was associated with augmentation of the humoral immune response to tumour antigens and removal of immune complexes. Again, our results suggest that SpA-induced interferon production may be implicated.

Thus, our results demonstrate that SpA induces IFN-γ production by nonadherent, T lymphocyte-depleted, FcR^+ human PBLs, and that IFN-γ augments both NK and ADCC activity. Our results suggest that the induction of IFN-γ by SpA exhibits several unique characteristics. Interferon induction by SpA seems to be less dependent on macrophages than induction of IFN-γ by other mitogens[14] and is apparently produced by non-T lymphocytes although more complete cell fractionation studies must be performed to confirm this preliminary observation. Maximum production of IFN-γ by SpA occurs within 24 h whereas other mitogens require 3–5 days for maximum IFN-γ production[14]. Finally, our data suggest that investigators using SpA as an immunological reagent for its Fc receptor-bearing properties must consider its interferon-producing potential in interpreting results.

We thank Joseph M. Davie, J. Russell Little Jr, Sondra Schlesinger and Gordon W. Philpott for critical reviews of this manuscript. We also thank Marsha Collett and Kimberly Merritt for help in preparing the manuscript. This work was supported in part by PHS grant 1 R26 CA23855, NCI grant 1 R26 CA28860 through the National Bladder Cancer Project, grant 1 R26 CA25792 from the National Prostate Cancer Project, Washington University Biomedical Research grant 3375-55874-B, General Research Support grant RIA-N50104-0 and Biomedical Research Support grant RIA-N90100-0 from The Jewish Hospital of St Louis, and American Cancer Society Institutional Research grant 1 N36Q. W.J.C. is recipient of American Cancer Society Junior Faculty Clinical Fellowship 382.

Received 8 April 1980; accepted 2 February 1981.

1. Bjork, I., Peterson, B.-A. & Sjoquist, J. *Eur. J. Biochem.* **29,** 579–584 (1972).
2. Kronvall, G. & Williams, R. C. *J. Immun.* **103,** 828–833 (1969).
3. Kronvall, G. & Gewurz, H. *Clin. exp. Immun.* **7,** 211–220 (1970).
4. Dossett, J. H., Kronvall, G., Williams, R. C. & Quie, P. G. *J. Immun.* **103,** 1405–1410 (1969).
5. Harvey, R. L., Kronvall, G., Troup, G. M., Anderson, R. F. & Williams, R. C. *Proc. Soc. exp. Biol. Med.* **135,** 453–456 (1970).
6. Forsgren, A., Svedjelund, A. & Wigzell, H. *Eur. J. Immun.* **6,** 207–213 (1976).
7. Sakane, T. & Green, I. *J. Immun.* **120,** 302–311 (1978).
8. Cowan, F. M., Klein, D. L., Armstrong, G. R. & Pearson, J. W. *Biomedicine* **30,** 23–27 (1979).
9. Sulica, A., Lakey, M., Gherman, M., Ghetie, V. & Sjoquist, J. *Scand. J. Immun.* **5,** 1191–1197 (1976).
10. Kay, H. D., Gonnard, G. D., West, W. H. & Herverman, R. B. *J. Immun.* **118,** 2058–2066 (1977).
11. Ray, P. K., Cooper, D. R., Bassett, J. G. & Mark, R. *Fedn Proc.* **38,** 1089 (1979).
12. Langford, M. P., Georgiades, J. A., Stanton, G. J., Dianzani, F. & Johnson, H. M. *Infect. Immun.* **26,** 36–41 (1979).
13. Catalona, W. J., Ratliff, T. L. & McCool, R. E. *J. Immun.* **122,** 1009–1014 (1979).
14. Epstein, L. B. *Tex. Rep. Biol. Med.* **35,** 42–56 (1977).
15. Wheelock, E. F. *Science* **169,** 310–311 (1965).
16. Herberman, R. B., Ortaldo, J. R. & Bonnard, G. D. *Nature* **277,** 221–223 (1979).
17. Vessella, R. L., Gormus, B. J., Lange, P. H. & Kaplan, M. E. *Int. J. Cancer* **21,** 594–603 (1978).
18. Pape, G. R., Troye, M., Axelsson, B. & Perlman, P. *J. Immun.* **122,** 2251–2260 (1979).

Production of immune interferon by murine T-cell clones from long-term cultures

Fabrizio Marcucci*, Michael Waller†, Holger Kirchner* & Peter Krammer†

* Institute of Virus Research and † Institute of Immunology and Genetics, German Cancer Research Centre, Heidelberg, FRG

Production of leukocyte interferon (IFN-α) and fibroblast interferon (IFN-β) can be induced by a variety of agents[1] but immune interferon, IFN-γ, is only obtained when lymphoid cells are stimulated by specific antigens, allo-antigens or T-cell mitogens[2,3]. Moreover, in bulk cultures, only small quantities of IFN-γ are produced. The type of cell producing IFN-γ has not been unambiguously defined[4–6] and so we set out to determine whether a pure T-cell population could produce it, exploiting the knowledge[7,8] that T cells can be maintained indefinitely in tissue culture by the addition of T-cell growth factors. Although not all T cells can found long-term cultures of this kind[9], cultures of antigen-specific helper, suppressor and killer T cells have been obtained[10–15] in this way. We now describe the production of substantial amounts of INF-γ when some (but not all) murine T-cell clones derived from such cultures are stimulated by either concanavalin A (Con A) or phytohaemagglutinin (PHA).

We obtained such a source of homogeneous T cells from clones of a 4-day culture of cells from the draining axillary and inguinal lymph nodes of AKR/J mice sensitized on the abdominal skin by 2,4,6-trinitrophenyl chloride (picryl chloride, PC) 5 days earlier[14]. These clones had been passaged in tissue culture for several months with a doubling time of ~20 h and grew strictly dependent on the presence of T-cell growth factors in the supernatant from rat spleen cells stimulated by Con A for 48 h. Cells from these clones were stimulated with Con A or PHA and tested for interferon production in different culture conditions.

The first series of experiments showed that these two mitogens induced interferon only when the T-cell growth factor-containing medium was replaced by fresh culture medium during interferon induction (Table 1). No antiviral activity could be detected without addition of T-cell mitogens or in the presence of B-cell mitogens (data not shown). The optimal dose for interferon induction, subsequently used for all further experiments was 5 or 10 μg ml^{-1} Con A or PHA. Table 1 indicates that interferon was also generated by mitogen stimulation in serum-free medium and its production significantly increased by addition of 2-mercaptoethanol to the medium. The antiviral activity in the supernatant of the stimulated T-cell clones was characterized as IFN-γ by the following criteria (outlined in Table 2): (1) species-restricted activity, (2) induction of the antiviral state dependent on intact RNA synthesis in the target cell, (3) sensitivity to pH 2 and to trypsinization, (4) resistance to RNase treatment, (5) persistence of the antiviral activity after ultracentrifugation, (6) loss of activity after incubation with anti-IFN-γ antiserum and persistence of activity after incubation with anti-IFN-α and β antiserum.

Thus, production of IFN-γ could be induced by T-cell mitogens from continuously growing, nontransformed T cells maintained by passage in long-term culture in medium containing T-cell growth factors and in the absence of adherent cells. Table 3 shows that this was true for several randomly selected T-cell clones. Furthermore, while most clones responded equally well to stimulation with Con A or PHA, some produced more interferon on stimulation with one or the other. Few clones gave only marginal responses; most produced IFN-γ at titres considerably higher than those obtained in bulk cultures stimulated with the same mitogens (data not shown). The capacity to produce IFN-γ in response to mitogen stimulation seemed to be a stable property of the clones as approximately the same titres were obtained from two randomly selected clones tested after a 2-month interval (Table 3).

Our cultures had not been established by deliberate isolation and further growth of single cells but by seeding 1 cell per culture well. As 5.2% of the initial culture wells were positive for growth, most of the established long-term cultures from these wells could be considered clones according to Poisson statistics. The clones originated from a bulk culture of more than 95% T cells and grew strictly dependent on the presence of T-cell growth factor-containing medium. Furthermore, IFN-γ release could only be stimulated by T-cell mitogens.

To characterize the phenotype of one of the clones producing IFN-γ we treated the clone PC AKR 96, from the same series as the other clones, with the monoclonal anti-Thy 1.1, anti-Lyt 1.2 or anti-Lyt 2.1 antibodies and complement. As shown in Table 4, the cells were resistant to lysis, evaluated by Trypan blue exclusion, even after two consecutive treatments with any of the above antibodies. This insensitivity to complement treatment for T-cell clones in long-term culture has been described by others[15]. Antibody-treated and control cells were resuspended in T-cell growth factor-containing medium and after 4 days, growth (evaluated by ^{3}H-thymidine uptake) and mitogen-induced IFN-γ production were determined. At this time, cells treated with anti-Thy 1.1+complement incorporated only low levels of ^{3}H-thymidine and almost no IFN-γ production could be observed. The phenomenon was specific because ^{3}H-thymidine uptake and IFN-γ production were unaltered after treatment of the cells with the monoclonal anti-Thy 1.2 antibody,

Table 1 Conditions for IFN-γ production by T-cell clones stimulated with Con A or PHA

			IFN-γ titres in different culture conditions (International Units per ml)			
Clone	Mitogen added	μg ml^{-1}	NCM	TCGFCCM	SF-NCM	NCM + 2-ME
PC AKR 29	None		<3	30	<3	10
	PHA	5	240	45	300	550
	PHA	10	550	30	550	720
	Con A	5	240	30	240	550
	Con A	10	720	45	720	1,000
PC AKR 54	None		<3	30	<3	5
	PHA	5	350	30	500	720
	PHA	10	240	45	240	500
	Con A	5	550	45	550	720
	Con A	10	300	45	300	450

NCM, normal culture medium, RPMI-1640 (Gibco), supplemented with L-glutamine (2 mM), streptomycin (100 μg ml^{-1}) + 5% fetal calf serum (FCS). TCGFCCM, T-cell growth factor-containing conditioned medium, as NCM with 2-mercaptoethanol (2-ME, 3×10^{-5}M) and substituted with 10 mM HEPES, 10% FCS and 10% supernatant containing T-cell growth factors. This supernatant was obtained by stimulation of 5×10^6 Sprague–Dawley rat spleen cells (70 ml in Falcon bottles (flat), 3024 F) with 5 μg Con A for 24 h. The stimulation was done in medium with the same constituents as above except that Iscove's modified Dulbecco's medium without HEPES was used. SF-NCM, serum-free normal culture medium. T-cell clones were obtained in the following way. The draining axillary and inguinal lymph node cells of AKR/J mice sensitized 5 days previously on the abdominal skin by PC were put into culture as described in ref. 14. On day 5 of culture, cells cleared from dead cells and debris by centrifugation over Ficoll–Urovison[14] were cloned by seeding 1 cell per well in 200 μl TCGFCCM containing 3×10^4 3,300-rad-irradiated AKR/J peritoneal exudate filler cells in 96-well flat-bottom microtitre plates (Linbro no. 76-003-05, Flow Laboratories). The old TCGFCCM was replaced once a week. After about 4 weeks all growing clones were separately transferred into 24-well plates (Linbro no. 76-033-05, Flow Laboratories) and after an additional 2 weeks the surviving clones were passaged and usually subcultured once a week in Falcon bottles (3024 F, Falcon) in TCGFCCM. All cultures were kept at 37° C, in humidified atmosphere, 95%–5% air–CO_2 mixture. For interferon induction, cells grown in TCGFCCM were centrifuged, washed twice and resuspended in NCM or in SF-NCM, or in NCM + 2-ME at a concentration of 1×10^6 ml^{-1} and stimulated with PHA or Con A. Culture supernatants were collected 24 h later and interferon was titrated by means of a plaque reduction assay using L cells and vesicular stomatitis virus as challenge virus[18].

which recognizes the allelic specificity of the Thy-1 locus not expressed on T cells from AKR mice. Typing of the Lyt markers for which monoclonal reagents were available (Lyt 1.2 and Lyt 2.1) was negative in all the assays performed. Thus, cells from this clone either are Lyt 1^-2^- or possess these surface antigens at low density and are therefore not detectable in our experimental conditions. Clones negative for the expression of these markers have already been described[15,16]. Bearing in mind that cultivation of long-term T-cell clones in T-cell growth factor-containing medium is accompanied by change in quantity or loss of expression of cell-surface antigens, particularly of the Ly series[15], these observations suggest that the cells from the clone PC AKR 96 are derived from a single T cell and continue to express the θ antigen.

The clones had been passaged in tissue culture for more than 3 months without addition of adherent cells before they were tested for IFN-γ production. We therefore assume that they were virtually free of macrophages and that IFN-γ was being produced by T-cell clones without help from adherent cells. That some clones yielded titres of IFN-γ considerably higher than those normally obtained from cell mixtures in bulk cultures stimulated with the same mitogens indicates that, in these conditions either we obtained an expansion of the IFN-γ-producer cells or some regulatory mechanism acting in bulk cultures was bypassed.

When cultured in the absence of growth factor-containing medium, T-cell clones produced IFN-γ after mitogen stimulation (Table 1). Unstimulated cells from these cultures did not

Table 2 Characterization of antiviral activity in the supernatant of murine T-cell clone PC AKR 29 stimulated with Con A or PHA

Treatment of supernatant	Activity* on mouse cells
pH 2 for 24 h	–
Centrifugation at 100,000g for 1 h	+
Trypsin (100 μg ml^{-1}) for 1 h at 37 °C	–
RNase (1000 μg ml^{-1}) for 1 h at 37 °C	+
Incubation with anti-mouse IFN-α and -β antiserum for 1 h at 37 °C	+
Incubation with anti-mouse IFN-γ antiserum for 1 h at 37 °C	–

Anti-mouse IFN-α and -β antiserum was a rabbit anti-L cell interferon antiserum with a neutralization titre of 1:6,000 on IFN-α and -β. Anti-mouse IFN-γ antiserum was a rabbit antiserum raised against PHA-induced, mouse spleen cell interferon with a neutralization titre of 1:640, ref. 19. +, Activity undiminished; −, 4>90% loss of activity.

* Loss of antiviral activity was observed after pretreatment of L cells with 1 μg ml^{-1} actinomycin D for 1 h. The antiviral activity was species restricted as no reduction of virus plaques was observed on Rita (monkey kidney) cells.

Table 3 IFN-γ production by different T-cell clones stimulated with Con A or PHA

	IFN-γ titres (IU ml^{-1}) after stimulation with			
	PHA		Con A	
T-cell clones	5 μg ml^{-1}	10 μg ml^{-1}	5 μg ml^{-1}	10 μg ml^{-1}
PC AKR 28	500	300	60	80
PC AKR 37	15	10	15	30
PC AKR 41	300	240	240	240
PC AKR 74	150	100	500	800
PC AKR 75	50	15	240	100
PC AKR 81	90	30	150	100
PC AKR 83	720	600	800	720
PC AKR 89	30	15	15	30
PC AKR 96	250	400	60	100
PC AKR 45/1*	240	800	30	80
PC AKR 45/2*	300	720	20	60
PC AKR 59/1*	720	800	720	720
PC AKR 59/2*	500	600	600	720

Stimulation of T-cell clones was done in NCM (see Table 1 legend).

* 1 and 2 represent the results from clones PC AKR 45 and 59 at a time interval of 2 months.

Table 4 Analysis of cell-surface markers of the T-cell clone PC AKR 96

	Characterization of the cells after treatment			
	Day 0		Day 4	
Treatment	% Living cells	IFN-γ production (IU ml^{-1})	Growth (^{3}H-TdR incorporation, c.p.m.)	IFN-γ production (IU ml^{-1})
Control	90	1,000	148,520	600
Complement (C′) alone	80	800	139,343	600
Anti-Thy 1.1+C′	75	720	3,047	10
Anti-Thy 1.2+C′	70	720	122,250	720
Anti-Lyt 1.2+C′	80	1,000	121,342	800
Anti-Lyt 2.1+C′	70	800	156,528	600

The control was NCM alone. Antibodies were monoclonal antibodies anti-Thy 1.1, anti-Lyt 1.2, Lyt 2.1 (code NEJ-014, NEJ-017 and NEJ-004, respectively, NEN) and anti-Thy 1.2 (code FfB5, Olac), and were used at a final dilution of 1:100, 1:50, 1:50 and 1:100, respectively. Complement was 1:6 diluted, selected rabbit complement. Treatment with antibodies and complement was performed twice with 2×10^6 cells as previously described[20]. At the end of the treatment the cells were resuspended in 1 ml NCM. 0.2 ml of each cell suspension was induced for IFN-γ production with PHA (10 μg ml^{-1}). 0.2 ml of a 1:10 dilution of each cell suspension were separately transferred into 24-well plates (Linbro no. 76-033-05, Flow Labs), each well containing 1 ml fresh TCGFCCM. 4 days later 3×0.2 ml of each cell suspension were separately transferred into a 96-well plate, pulsed for 4 h with 1 μCi ml^{-1} of tritiated thymidine (^{3}H-TdR, Amersham) after which the cultures were aspirated on filter paper by use of a semi-automated collecting device (Skatron, Flow Labs) and radioactivity of the filters measured in a scintillation counter (Mark II, Searle-Nuclear, Chicago). Another 0.2-ml aliquot was centrifuged and resuspended in 0.2 ml NCM containing PHA (10 μg ml^{-1}). 24 h later the supernatants were tested for IFN-γ as described in Table 1 legend. Per cent living cells was determined by Trypan blue exclusion. Growth results are the mean value of triplicate samples.

release antiviral activity. Supernatants from unstimulated cells cultured with T-cell growth factor-containing medium were positive for IFN-γ because the medium itself contained IFN-γ. Addition of mitogens to cells from these cultures did not increase IFN-γ titres above the medium background (Table 1). This might have been due to blocking activity exerted by IFN-γ similar to the blocking observed for IFN-α and β production[17].

T-cell help, cytotoxic activity and suppression can be obtained from T-cell lines and clones without antigen or mitogen stimulation[10–12]. Our data indicate that other T-cell activities, like IFN-γ production, can be stimulated from T-cell lines and clones only with mitogens. It remains to be determined whether other stimuli, such as specific antigens, which resemble more closely the *in vivo* situation, are also able to induce IFN-γ production from T-cell clones.

Finally, the fact that large quantities of IFN-γ can be produced in serum-free conditions (Table 1) from homogeneous cells in continuous culture could be useful for its biochemical isolation.

We thank Professor K. Resch and Drs S. Landolfo and M. Simon for helpful discussions, Ms A. Michnay and M. Nowak for technical assistance and Ms M. Kasamasch for typing the manuscript. Drs June Dunnick and Juana Wietzerbin provided the anti-mouse IFN-α and -β antiserum and anti-mouse IFN-γ antiserum respectively.

Received 13 October 1980; accepted 25 February 1981.

1. Stewart, W. E. II in *The Interferon System*, 25–27 (Springer, New York, 1979).
2. Gifford, G. E., Tibor, A. & Peavy, D. L. *Infect. Immun.* **3,** 164–166 (1971).
3. Wheelock, E. F. *Science* **149,** 310–311 (1965).
4. Sonnenfeld, G., Mandel, A. D. & Merigan, T. C. *Immunology* **36,** 883–890 (1979).
5. Neumann, C. & Sorg, C. *Eur. J. Immun.* **8,** 582–589 (1978).
6. Kirchner, H. *et al. Eur. J. Immun.* **9,** 824–826 (1979).
7. Gillis, S. & Smith, K. A. *Nature* **268,** 154–156 (1977).
8. Morgan, D. A., Ruscetti, F. W. & Gallo, R. C. *Science* **193,** 1007–1008 (1976).
9. V. Boehmer, H., Haas, W., Pohlit, H., Hengartner, H. & Nabholz, M. *Springer Semin. Immunopath.* **3/1,** 23–27 (1980).
10. Watson, J. D. *J. exp. Med.* **150,** 1510–1519 (1979).
11. McVay-Boudreau, L., Fresno, M., Nabel, G. & Cantor, H. *Abstr. 4th int. Congr. Immun.*, Paris, 4.2.14 (1980).
12. Baker, P. E., Gillis, J. & Smith, K. A. *J. exp. Med.* **149,** 273–278 (1979).
13. Nabholz, M., Engers, H. D., Collavo, D. & North, M. *Curr. Topics Microbiol.* **81,** 176–187 (1978).
14. Krammer, P. H., Rehberger, R. & Eichmann, K. *J. exp. Med.* **151,** 1166–1182 (1980).
15. Nabholz, M. *et al. Immun. Rev.* **51,** 125–156 (1980).
16. Glasebrook, A. L. & Fitch, F. W. *J. exp. Med.* **151,** 876–895 (1980).
17. Stewart, W. E. II, Grosser, L. B. & Lockart, R. Z. Jr *J. gen. Virol.* **13,** 35–50 (1971).
18. Campbell, J. B., Gruberger, T., Kochmann, M. A. & White, S. L. *Can. J. Microbiol.* **3,** 1247–1253 (1975).
19. Stefanos, S., Catinot, L., Wietzerbin, J. & Falcoff, E. *J. gen. Virol.* **50,** 225–229 (1980).
20. Simon, M. M. & Abenhardt, B. *Eur. J. Immun.* **10,** 334–341 (1980).

Herpesvirus-transformed cytotoxic T-cell lines

Donald R. Johnson & Mikael Jondal

Department of Tumor Biology, Karolinska Institutet,
S-104 01 Stockholm, Sweden

Investigations of cellular cytotoxicity of the immune system are hampered by the lack of continuously growing, transformed cell lines which express a cytotoxic potential. Here we describe cytotoxic cell lines from the cotton-topped marmoset monkey, transformed by Herpesvirus Ateles (HVA) or Herpesvirus Saimiri (HVS), which can kill certain target cells in a short-term *in vitro* test. HVA/HVS-transformed cells have earlier been classified as belonging to the T-cell lineage[1–3] in contrast to Epstein–Barr virus (EBV)-transformed cells derived from B lymphocytes[4]. We suggest that the HVA/HVS-transformed killer cell lines described here represent an effector population resembling, or corresponding to, marmoset natural killer (NK) cells and that they may be used to define the cytolytic mechanism involved in cellular cytotoxicity and possibly also effector cell receptors and target cell antigens, as well as regulatory mechanisms of general biological interest.

Peripheral blood lymphocytes from adult cotton-topped marmosets were separated on Ficoll–Isopaque and transformed by cell-free supernatants from HVA-infected marmoset kidney fibroblasts[3]. The resulting cell lines were propagated as suspension cultures in RPMI 1640 media supplemented with 10% fetal calf serum and antibiotics. A 6-h ^{51}Cr release cytotoxicity assay was performed in microplates as described in Table 1 legend, using various target cells. Cytotoxicity assays in which HVA-carrying cell lines were used as effector cells, were compared with cytotoxicity expressed by marmoset and human NK cells isolated from the peripheral blood of normal healthy donors. Five different HVA-transformed cell lines and two HVS-transformed cell lines were tested for cytotoxic activity against the human T-cell line Molt-4 (Table 1) derived from acute lymphocytic leukaemia[5]. All seven HVA/HVS-transformed cell lines, and human NK cells, were cytotoxic in contrast to human fetal thymocytes and marmoset B-cell lines. Both *in vitro*-transformed cell lines and lines transformed *in vivo*, that is, tumour-derived cells, were cytotoxic, which indicates that the cytotoxic

0028-0836/81/190081—03$01.00

Table 1 Cytotoxicity by HVA/HVS-transformed cell lines from cotton-topped marmoset monkeys

Effector cell	Transforming virus	Origin	Type	Age	Cytotoxicity against target cell Molt-4 (specific ^{51}Cr release)		
					25:1	12.5:1	6.3:1
Experiment I							
77-DI-2	HVA	*In vitro* transformed	T	6 months	68.4	59.9	48.9
70-KI-1	HVA	Same	T	3 yr	60.1	49.6	38.2
22-CM-37	HVA	Same	T	4 yr	11.4	9.8	7.9
22-CM-61	HVA	Same	T	4 yr	32.2	29.6	19.8
1022	HVA	Tumour-derived	T	6 yr	62.3	52.1	44.8
70-N-2	HVS	Same	T	8 yr	52.1	40.8	33.8
MLC	HVS	Same	T	9 yr	19.2	15.3	10.7
Human NK	—	Peripheral blood	—	Fresh	68.7	48.3	28.4
Human thymocytes	—	Fetal thymus	—	Fresh	0.2	−0.1	0.0
Experiment II							
77-DI-2	HVA	*In vitro* transformed	T		66.2	48.7	39.4
77-DI-2	EBV	Same	B		3.1	0.6	0.1
70-KI-1	EBV	Same	B		0.2	0.1	0.2
Human NK	—	Peripheral blood	—		75.4	63.6	54.1

Cytotoxicity assay was performed in V-shaped microtitre plates using 10^4 ^{51}Cr-labelled target cells. RPMI 1640 with 10% fetal calf serum was used as testing medium, giving a total volume of 150 μl per well. After 6 h, 50 μl was collected from each well and cytotoxicity calculated as follows: test release − spontaneous release/80% of maximum label − spontaneous release × 100. Values are means of duplicate wells; two representative experiments are shown.

potential of these cell lines is a stable characteristic that does not change even after many years in tissue culture (note particularly the old cell lines 1022, MLC and 70-N-2).

As the Molt-4 cell is a standard target for human NK cells[6], and as the HVA/HVS cells are spontaneously cytotoxic (that is, no special triggering is needed), we investigated whether they would have the same target specificity as human NK cells when tested against an extended panel of cells (see Table 2). HVA/HVS cell lines were found to kill most efficiently target cells derived from T-cell malignancies, such targets are also readily killed by human NK cells[7]. The marmoset lines killed K562 but to a lesser degree than human NK cells (K562 is also a standard target for the human NK system[6,8]). HVA/HVS cell lines killed B-cell lines from non-EBV (as opposed to EBV-containing) malignancies. In contrast, human NK cells seem to kill malignant as opposed to non-malignant cells, as previously shown[9]. Neither HVA/HVS lines nor human NK cells kill normal mitogen-transformed T cells—this was also confirmed using T blasts from concanavalin A and MLC cultures (data not shown). Finally, the HVA/HVS cell lines did not kill the standard NK target in the mouse system, YAC-1, a Moloney leukaemogenic virus (MLV)-transformed cell line of thymic origin[8,10].

Comparison of NK cells from cotton-topped marmosets with human NK cells (Table 3) showed that marmoset NK cytotoxicity resembled that of the 77-DI-2 cell line as both predominantly killed the Molt-4 target and, to a much lesser degree, the Raji and K-562 targets. In contrast, human NK cells demonstrated highest activity against the K-562 target. These results indicate that the difference in target cell specificity between the 77-DI-2 cell line and human NK cells (Table 2) reflects minor differences between the human and marmoset NK systems.

To ascertain that the cell line-mediated cytotoxicity is a true membrane-mediated event (as are conventional T- and NK-cell killing) and not dependent on secreted lymphotoxin-like molecules, we investigated cytotoxicity on the single cell level using a recently described technique in which separate effector–target cell conjugates were embedded in agarose[11]. We also compared the number of active killer cells in the 77-DI-2 cell line with those present in normal marmoset lymphocytes propagated by T-cell growth factor (TCGF). TCGF was prepared from supernatants of phytohaemagglutinin (PHA)-stimulated human lymphocytes, usually pooled from three different donors, as described elsewhere[12]. Table 4 shows that 38.5% of the HVA-transformed killer cells were able to bind to the target cells but

Table 2 Target cell specificity of the 77-DI-2 effector cell line

Target cell	Origin	Characteristics	EBV	Effector cells (specific ^{51}Cr release)	
				77-DI-2	Human NK
Molt-4	Acute lymphocytic leukaemia	T cell	−	68.9	68.7
1301	Same	Same	−	49.3	66.4
CCRF-CEM	Same	Same	−	71.2	68.9
HSB-2	Same	Same	−	71.6	79.7
HD-Mar	Hodgkin's disease	Same	−	36.5	94.8
BJAB	Burkitt's lymphoma	B cell	−	29.1	94.0
Ramos	Same	Same	−	15.9	20.6
K562	Chronic myelocytic leukaemia	CML cell	−	17.3	87.5
DHL-9	Diffuse histiocytic lymphoma	B cell	−	29.4	52.8
Daudi	Burkitt's lymphoma	Same	+	−0.3	42.8
P3HR-1	Same	Same	+	0.1	65.4
B95-8	Normal marmoset B cells	Same	+	6.3	41.8
LCL(IE)	Normal human peripheral blood	Same	+	0.4	1.2
Human T-cell blasts	Normal human T cells stimulated by PHA	T cell	−	0.2	1.6
YAC-1	Mouse line, MLV-transformed	Same	−	0.7	9.9

Cytotoxicity assay was performed as described in Table 1 legend. The data represent three different experiments performed in identical conditions using an effector/target ratio of 25:1.

only 18.3% of these were active killer cells. These values were comparable with those found for the two TCGF-maintained cultures, which demonstrates equivalent cytotoxic potential for these two effector cell types.

TCGF has been shown to sustain proliferation of antigen-specific T killer[13] cells but also to generate NK-like activity[14], and it has been suggested that specific killing and NK killing may be expressed simultaneoulsy in activated T cells at certain stages in T-cell differentiation[15]. Thus, because of uncertainties about the true nature of TCGF cultures and their instability, it seems that HVA/HVS-transformed cytotoxic cell lines are useful, as viral transformation probably freezes effector cells in a defined stage of differentiation representative of the cytotoxic potential expressed by the corresponding non-transformed cells. An analogy could be drawn here between HVA/HVS and EBV because EBV only transforms B lymphocytes in a defined stage of their differentiation[4].

We found marked variation in the number of active killer cells in the 77-DI-2 cell line when tested at different time points (data not shown), the results in Table 4 being of average activity. The explanation for these variations is unclear, suggesting that regulatory mechanisms occur during tissue culture, but we know only that fast growing cultures maintained in roller bottles show the highest activity. The possibility of artificial killing by mycoplasma was excluded by comparing the activity of 77-DI-2 cells kept in mycoplasma-free conditions with that of cells in ordinary tissue culture—similar cytotoxic activities were recorded (data not shown).

Table 3 Natural cytotoxicity by peripheral blood lymphocytes from cotton-topped marmoset monkeys

	Cytotoxicity against target cell (specific ^{51}Cr release)		
Effector cells	1301	K-562	Raji
Monkey 1	43.4	12.1	11.2
Monkey 2	53.8	9.2	10.6
Monkey 3	33.9	8.2	6.2
77-DI-2 cell line	56.2	5.6	4.1
Human NK cells	31.5	72.3	9.3

Cotton-topped marmoset peripheral blood was obtained from Dr L. Wolfe of the Rush Medical Center, Chicago, USA. Lymphocytes were isolated on arrival and tissue-cultured overnight before testing. The cytotoxicity assay was performed using an effector/target cell ratio of 10:1 as described in Table 1 legend. Similar results were obtained in three separate repeat experiments.

Table 4 Cytotoxicity by the 77-DI-2 effector cell line and by TCGF-propagated normal marmoset lymphocytes

	Cytotoxicity (specific ^{51}Cr release)				
	3.0:1	1.5:1	0.8:1	TBCs* (%)	Cytotoxic TBCs† (%)
77-DI-2 (HVA)	50.2	40.7	31.6	38.5	18.3
77-DI-2 (TCGF)	58.6	52.5	50.7	28.7	23.1
77-EC-1 (TCGF)	72.9	65.6	56.2	31.8	30.4

Cytotoxicity was performed as described in Table 1 legend. Molt-4 was used as the target cell.

* Target binding cells (TBCs) were assayed as described elsewhere[17,18].

† Cytotoxic TBCs were tested by the agarose single cell cytotoxicity assay[11]. Briefly, effector and target cells were conjugated at a ratio of 1:10 using carboxyfluorescein-diacetate-labelled target cells, and layered in agarose on to plastic Petri dishes. TBCs were estimated at 0 h and cytotoxicity at 6 h.

Morphological analysis of the 77-DI-2 cell line revealed that a substantial part of the cells resembled large granular lymphocytes (LGLs) and that this fraction seemed to be conjugated to susceptible target cells (data not shown). This may be significant as human NK cells have been reported as having LGL morphology[8]. Electron microscopy of 77-DI-2 cells showed immature cells with abundant endoplasmic reticulum and uropod formation. This indicates a secretory function and a high degree of mobility (data not shown).

Results from electron microscopy and the high cytotoxic activity on some target cells in ^{51}Cr release assays compared with the number of active killer cells found in the agarose single cell cytotoxicity assay, led us to ask the question whether 77-DI-2 effector cells were able to recycle, that is, to kill more than one target cell in the given test period. We previously found this to be a characteristic of interferon-activated human NK cells[16,17]. Using the same technique as for the latter to demonstrate recycling, we found a marked recycling capacity in the 77-DI-2 cell line (data not shown). This recycling capacity is also characteristic of immune cytotoxic T cells[18].

In conclusion, we have found that HVA/HVS-transformed T-cell lines from cotton-topped marmosets can kill predominantly human T-cell lines and non-EBV B-cell lines of malignant origin. This reactivity is clearly a cell-mediated event requiring cell contact and not caused by lymphotoxin-like molecules. The specificity of HVA/HVS cell line killing resembles that of the marmoset NK system, and may be due to these viruses transforming marmoset NK cells or a subpopulation of them. If so, the low reactivity against EBV-induced tumours could indicate that marmoset NK cells differ from human NK cells in their target cell specificity. The selective EBV target deficiency of the marmoset NK system could possibly explain the high degree of susceptibility to EBV-induced lymphoproliferation/lymphomas in these animals[4]. We are now investigating this question.

Alternatively, the HVA/HVS cell line killing could be unrelated to the NK system and represent either a hitherto undefined effector population or possibly a T-effector population which is continuously generated in an autoreactive MLC type of culture. In either case, we consider that the HVA/HVS-transformed cytotoxic cell lines described here are the first immortalized cell lines of lymphoid origin reported which have the capacity to kill by cell contact, without added cell agglutinants, and which display a defined specificity. Possibly they represent an immune function of biological significance and will be useful for studying the finer mechanisms of cell-mediated cytotoxicity reactions.

We thank Maj-Britt Alter for technical assistance, Dr L. Wolfe for a continuous supply of marmoset blood and Dr L. Falk for supplying several of the investigated HVA/HVS cell lines. We also thank Dr G. Klein for discussing the data and manuscript. This work was supported by funds from the Swedish Cancer Society, Karolinska Institute and Karolinska Hospital and by contract NO1-CP-33316 from the Division of Cancer Cause and Prevention, NCI. D. R. J. is a recipient of a fellowship from the Cancer Research Institute, New York.

Received 25 November 1980; accepted 9 March 1981.

1. Deinhardt, F. & Deinhardt, J. in *The Epstein-Barr Virus* (eds Epstein, M. A. & Achong, B. G.) 373–415 (Springer, New York, 1979).
2. Falk, L., Wright, J., Wolfe, L. & Deinhardt, F. *Int. J. Cancer* **14,** 244–251 (1974).
3. Falk, L., Johnson, D. R. & Deinhardt, F. *Int. J. Cancer* **21,** 652–657 (1978).
4. Miller, G. in *The Epstein-Barr Virus* (eds Epstein, M. A. & Achong, B. G.) 351–372 (Springer, New York, 1979).
5. Minowada, J., Ohnuma, T. & Moore, G. E. *J. natn. Cancer Inst.* **49,** 891–895 (1972).
6. Jondal, M. & Pross, H. *Int. J. Cancer* **15,** 596–605 (1975).
7. Ono, A., Amos, D. B. & Koren, H. S. *Nature* **266,** 546–548 (1977).
8. Möller, G. (ed.) *Immun. Rev.* **44** (1979).
9. Jondal, M., Spina, C. & Targan, S. *Nature* **272,** 62–64 (1978).
10. Kiessling, R., Klein, E. & Wigzell, H. *Eur. J. Immun.* **5,** 112–117 (1975).
11. Grimm, E. A., Thoma, J. A. & Bonavida, B. *J. Immun.* **123,** 2870–2877 (1979).
12. Bonnard, G. D., Ysaka, K. & Maca, R. D. *Cell. Immun.* **51,** 390–401 (1980).
13. Möller, G. (ed.) *Immun. Rev.* **51** (1980).
14. Massuci, M. G., Klein, E. & Argov, S. *J. Immun.* **124,** 2458–2463 (1980).
15. Klein, E. *Immun. Today* **1,** 4–6 (1980).
16. Ullberg, M. & Jondal, M. *J. exp. Med.* **153,** 615–628 (1981).
17. Ullberg, M., Merrill, J. & Jondal, M. *Scand. J. Immun.* (in the press).
18. Berke, G. *Prog. Allergy* **27,** 69–116 (1979).

N-acetoxy-acetylaminofluorene reacts preferentially with a control region of intracellular SV40 chromosome

Peter Beard*, Motohisa Kaneko† & Peter Cerutti†

Departments of *Virology and †Carcinogenesis, Swiss Institute for Experimental Cancer Research, 1066 Epalinges, Switzerland

Many chemical carcinogens or their metabolites react with DNA[1,2]; thus it is of interest to determine what effect chromosomal structure has on these reactions. The chromosome of simian virus 40 (SV40) is well suited for such studies; like chromatin of eukaryotic cells, it is organized into nucleosomes. The nucleotide sequence of SV40 is known[3,4], together with much about the pattern of viral gene expression and DNA replication[5], and the structure of the viral chromosome[6–14]. We have investigated the binding of the ultimate carcinogen, N-acetoxy-acetylaminofluorene (AAAF), to specific regions of the SV40 chromosome *in situ* in the intact infected cell. The results, reported here, indicate that a region containing regulatory functions on the intracellular SV40 chromosome has unique structural properties which render it more susceptible to attack by AAAF than the rest of the SV40 genome. The preferential binding of AAAF to regulatory regions of chromatin may have implications for the mechanism of action of this and similar carcinogens.

AAAF binds covalently to DNA both *in vitro* and in cells growing in culture[15–18]. Over 95% of the reaction products involve guanine residues[15,16,18]. Like other carcinogens which introduce bulky substituents to DNA bases, AAAF reacts more extensively with internucleosomal linker DNA than with nucleosomal core DNA[18–22].

We infected monolayer cultures of the monkey kidney cell line BSC-B with SV40. In the late phase of the lytic cycle, during viral DNA replication, the medium was replaced with medium containing ^{3}H-AAAF (see Table 1 for details). After labelling for 15 min at 37 °C the cells were washed and SV40 DNA selectively extracted[23] and purified by centrifugation in gradients of CsCl–ethidium bromide[24]. As a control we treated protein-free SV40 DNA *in vitro* with ^{3}H-AAAF in the following way. SV40 DNA was extracted from infected cells by the method described above. The supercoiled DNA was then treated with DNA topoisomerase[25] to relax superhelical turns.

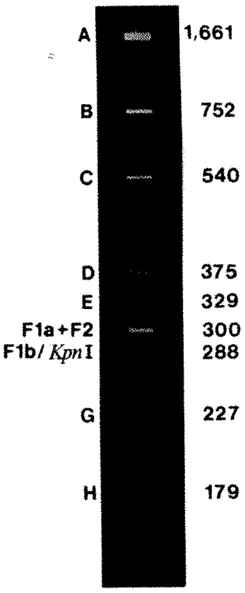


Fig. 1 Low-gelling-temperature agarose gel electrophoresis of SV40 DNA after digestion with restriction enzymes *Hae*III plus *Kpn*I. DNA was electrophoresed in a 3% gel of low-gelling-temperature agarose (Sigma) in Tris-phosphate–EDTA buffer, *p*H 7.8, for 6 h at 4 V cm^{-1}. The gel was stained with ethidium bromide (1 µg ml^{-1}) and photographed in light of 254 nm wavelength. Fragments resulting from cleavage by *Hae*III alone are named according to Lebowitz *et al.*[38]. The larger fragment produced by *Kpn*I cleavage of *Hae*III fragment F1b is marked *Hae*III F1b/*Kpn*I; the smaller product of this cleavage is too short to be detected in this gel (see Fig. 2*c*). Fragment lengths in base pairs are from ref. 5.

Table 1 Distribution of radioactivity in SV40 DNA restriction fragments after labelling of DNA *in vitro* or intracellularly with ^{3}H-AAAF

SV40 *Hae*III fragments	Per cent of total guanines in each fragment	Per cent of total radioactivity in each fragment: Labelled as free DNA	Per cent of total radioactivity in each fragment: Labelled *in situ*
A	30	29	27
B	15.9	18.1	16.1
C	11.1	11.6	9.7
D	7.4	6.5	8.6
E	7.4	6.5	6.6
F1a+F2	12.0	12.3	12.7
F1b/*Kpn*I	7.2	7.6	11.5
G	4.9	4.6	4.4
H	3.9	3.5	3.3

Intracellular SV40 DNA was labelled by addition of ^{3}H-AAAF (9 µM, 878 mCi mmol^{-1}, Midwest Research Institute, Kansas City) to virus-infected cell cultures during 15 min at 40 h after infection. SV40 DNA was then extracted and purified[23,24]. Free DNA (100 µg ml^{-1}) was labelled in solution in 10 mM Tris-HCl, 1 mM EDTA *p*H 7.5. Unincorporated AAAF was then removed by repeated extractions with ether and precipitation of the DNA with ethanol. Labelled DNA (2–10 × 10^3 c.p.m.) was digested with endonucleases *Hae*III plus *Kpn*I (New England Biolabs) and electrophoresed in preparative gel in conditions described in Fig. 1 legend. To determine the radioactivity in each restriction fragment, a simple and reliable procedure was developed. Slices of low-gelling-temperature agarose containing the DNA bands were cut, and water added to 1 ml. The agarose was melted at 70 °C for 5 min and 15 ml Aquasol (NEN) were added and the suspension mixed briefly until homogeneous. Samples were cooled in the dark for several hours, then counted in a scintillation spectrometer. Gel slices with no DNA gave the same background values as Aquasol alone (30 c.p.m.). The radioactivity in each fragment after subtraction of the background is given as a percentage of the total radioactivity. The percentage of total guanines in each fragment was determined from the SV40 nucleotide sequence[3,4].

Relaxed circular DNA is a more appropriate control than supercoiled DNA as SV40 chromosomes contain topoisomerase activity and are themselves thought not to be under torsional stress[26,27]. The free DNA was reacted with ^{3}H-AAAF as described in Table 1. The adduct concentration in SV40 DNA which had been exposed to the carcinogen *in vitro* or *in situ* in the infected cell was estimated from the specific activity of the ^{3}H-AAAF and the radioactivity content of the DNA. SV40 DNA labelled *in vitro* contained an average of about 10 adducts per molecule, whereas SV40 DNA labelled *in situ* contained on average 2 adducts per molecule.

AAAF-modified SV40 DNA was digested with a mixture of the restriction endonucleases *Hae*III and *Kpn*I. This combination gives a set of fragments of convenient size which can be separated by electrophoresis in gels of 3% low-gelling-temperature agarose (Fig. 1). In no case was the AAAF-modified DNA seen to give an altered restriction enzyme digestion pattern compared with normal SV40 DNA. Table 1 shows, for each fragment, its proportions of the total guanine content (calculated from the nucleotide sequence[3,4]) and of the incorporated radioactivity after *in vivo* or *in vitro* labelling. The amount of radioactivity in fragments labelled as free DNA followed the guanine content; when the specific activity was determined as a function of guanine content (Fig. 2*b*), all fragments were labelled equally within experimental error. This result is expected if no region of SV40 DNA shows preferential binding of AAAF because over 95% of AAAF–DNA adducts are formed with guanine residues[15,16,18]. In contrast, the results obtained with SV40 DNA isolated from ^{3}H-AAAF-treated cells indicated non-random distribution of AAAF–guanine adducts. In four determinations the fragment *Hae*III F1b/*Kpn*I was always labelled to a specific activity 1.5–2 times higher than that expected from the guanine content (Fig. 2*a*). The adduct concentrations in the restriction fragments from the rest of the SV40 genome were proportional to their guanine content.

0028-0836/81/190084—02$01.00

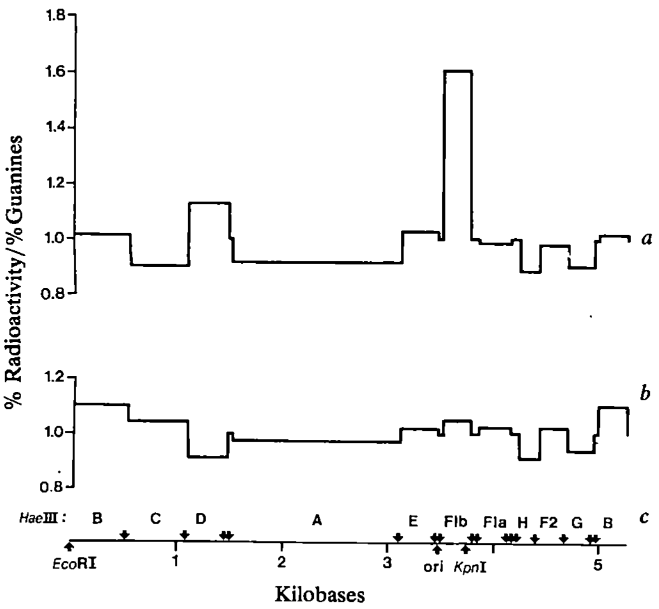


Fig. 2 Preferred reaction site for AAAF on the intracellular SV40 chromosome. The percentage of total guanines present in each SV40 restriction fragment, and the percentage of total ^{3}H-AAAF radioactivity in each fragment were determined as described in Table 1 legend. The ratio of the percentage of total radioactivity to the percentage of total guanines was calculated for each fragment. *a*, SV40 DNA labelled *in situ* in infected cells; *b*, labelled as free SV40 DNA; *c*, the abcissa is a *Hae*III endonuclease cleavage map of the circular SV40 DNA interrupted at the *Eco*RI site. The map is based on Lebowitz *et al.*[38], as is the SV40 nucleotide sequence[3,4]. DNA fragments too small to be analysed have been assumed to be labelled at the average value for the whole DNA.

Thus in the intracellular SV40 chromosome, a stretch of DNA within ~300 base pairs of the 'late genes' side of the origin of replication (Fig. 2*c*) is preferentially-reactive with AAAF. We do not know whether the region just to the 'early genes' side of the origin is also specifically overlabelled with AAAF, as the restriction fragment containing this region is too small (43 base pairs) to be analysed in these experiments. The next distal fragment to the early side, *Hae*III E (329 base pairs) was labelled normally.

The level found for AAAF modification of the hyper-reactive region is a minimum estimate because we do not know the exact limits of this region on the SV40 genome. If the exposed region is much shorter than the DNA fragment *Hae*III F1b/*Kpn*I which we characterized as containing it, then the level of modification would be correspondingly higher than the value we obtained, which is the average for the whole of that fragment. The level would also be greater if the hyper-reactive region were present in only a fraction of the SV40 chromosomes in the cell.

Why should AAAF react predominantly with a certain region of chromosomal DNA? The higher reactivity of AAAF towards nucleosomal linker compared with core DNA[18–22] suggests that the hyper-reactive region of SV40 chromosomes may represent DNA free of protein, or associated with less tightly bound protein. Electron microscopic observation[11,12] showed a nucleosome-free region on a fraction of isolated SV40 chromosomes; it was unclear to what extent this region already existed *in vivo* or was created by loss of nucleosomes during isolation of nuclei or of the viral chromosomes[11,12]. From our results, it is likely that a nucleosome-free stretch of DNA exists on SV40 chromosomes also *in situ* in the infected cell. The reaction of AAAF with DNA is also known to favour denatured DNA over native[28]; it is therefore possible that the hyper-reactive region has an unusual secondary or tertiary structure (reviewed in ref. 29). This region is not, however, simply the most readily denatured part of SV40 as detected by electron microscopy or S_1 nuclease sensitivity[30,31]. It also remains possible that AAAF induces an alteration of chromosomal structure in the SV40 origin region. However, modification of DNA by AAAF has been reported not to interfere with the *in vitro* reconstitution of apparently normal nucleosomes[32].

The region of the SV40 chromosome which reacts preferentially with AAAF contains regulatory sequences controlling the origin of replication of DNA[33], T-antigen binding sites[34] and likely promoters for late transcription[35,36]. SV40 uses cellular enzymes for its replication[5], and SV40 chromatin probably resembles certain regions of cellular chromatin in structure[37]. Therefore, regulatory sequences of some cellular genes may likewise be particularly susceptible to modification by AAAF and similar carcinogens.

We thank B. Hirt for advice and criticism of the manuscript, and K. Nyfeler and B. Bentele for technical assistance. This work was supported by grants 3.299.78 and 3.305.78 from the Fonds National Suisse de la Recherche Scientifique.

Received 9 December 1980; accepted 9 March 1981.

1. Sarma, D. S. R., Rajalakshmi, S. & Farber, E. in *Cancer* Vol. 1 (ed. Becker, F. F.) 235–287 (Plenum, New York, 1975).
2. Neidle, S. *Nature* **283**, 135 (1980).
3. Fiers, W. *et al. Nature* **273**, 113–120 (1978).
4. Reddy, V. B. *et al. Science* **200**, 494–502 (1978).
5. Tooze, J. (ed.) *DNA Tumor Viruses* (Cold Spring Harbor Laboratory, New York, 1980).
6. Bellard, M., Oudet, P., Germond, J.-E. & Chambon, P. *Eur. J. Biochem.* **70**, 543–553 (1976).
7. Cremisi, C., Pignatti, P. & Yaniv, M. *Biochem. biophys. Res. Commun.* **73**, 548–554 (1976).
8. Ponder, B. A. J. & Crawford, L. V. *Cell* **11**, 35–49 (1977).
9. Scott, W. A. & Wigmore, D. J. *Cell* **15**, 1511–1518 (1978).
10. Varshavsky, A. J., Sundin, O. & Bohn, M. *Cell* **16**, 453–466 (1979).
11. Saragosti, S., Moyne, G. & Yaniv, M. *Cell* **20**, 65–75 (1980).
12. Jakobovits, E., Bratosin, S. & Aloni, Y. *Nature* **285**, 263–265 (1980).
13. Beard, P. *Cell* **15**, 955–967 (1978).
14. Mathis, D. J., Oudet, P., Wasylyk, B. & Chambon, P. *Nucleic Acids Res.* **5**, 3523–3547 (1978).
15. Kriek, E. *Biochim. biophys. Acta* **355**, 177–203 (1974).
16. Kriek, E., Miller, J. A., Juhl, U. & Miller, E. C. *Biochemistry* **6**, 177–182 (1967).
17. Amacher, D. E. & Lieberman, M. W. *Biochem. biophys. Res. Commun.* **74**, 285–290 (1977).
18. Kaneko, M. & Cerutti, P. A. *Cancer Res.* (in the press).
19. Cech, T. & Pardue, M. *Cell* **11**, 631–640 (1977).
20. Wiesehahn, G., Hyde, J. & Hearst, J. *Biochemistry* **16**, 925–932 (1977).
21. Jahn, C. & Litman, G. *Biochemistry* **18**, 1442–1449 (1979).
22. Metzger, G., Wilhelm, F. & Wilhelm, M. *Biochem. biophys. Res. Commun.* **75**, 703–710 (1977).
23. Hirt, B. *J. molec. Biol.* **26**, 365–369 (1967).
24. Radloff, F., Bauer, W. & Vinograd, J. *Proc. natn. Acad. Sci. U.S.A.* **57**, 1514–1521 (1967).
25. Champoux, J. J. & McConaughy, B. L. *Biochemistry* **15**, 4638–4642 (1976).
26. Sen, A. & Levine, A. J. *Nature* **249**, 343–344 (1974).
27. Young, L. S. & Champoux, J. J. *Nucleic Acids Res.* **5**, 623–635 (1978).
28. Levine, A. F., Fink, L. M., Weinstein, I. B. & Grunberger, D. *Cancer Res.* **34**, 319–327 (1974).
29. Davies, D. R. & Zimmerman, S. *Nature* **283**, 11–12 (1980).
30. Beard, P., Morrow, J. F. & Berg, P. *J. Virol.* **12**, 1303–1313 (1973).
31. Mulder, C. & Delius, H. *Proc. natn. Acad. Sci. U.S.A.* **69**, 3215–3219 (1972).
32. Yamasaki, H., Leffler, S. & Weinstein, I. B. *Cancer Res.* **37**, 684–691 (1977).
33. Dana, K. J. & Nathans, D. *Proc. natn. Acad. Sci. U.S.A.* **69**, 3097–3100 (1972).
34. Tjian, R. *Cell* **13**, 165–179 (1978).
35. Ghosh, P. K., Reddy, V. B., Swinscoe, J., Lebowitz, P. & Weissman, S. *J. molec. Biol.* **126**, 813–846 (1978).
36. Handa, H., Kaufman, R. J., Manley, J., Gefter, M. & Sharp, P. A. *J. biol. Chem.* **256**, 478–482 (1981).
37. Shelton, E. R., Wassarman, P. M. & DePamphilis, M. L. *J. molec. Biol.* **125**, 491–514 (1978).
38. Lebowitz, P., Siegel, W. & Sklar, J. *J. molec. Biol.* **88**, 105–123 (1974).

Location of domains in globular proteins

Alexander A. Rashin

Institute of Biological Physics, Academy of Sciences of the USSR,
142292 Pushchino, Moscow Region, USSR

Although it has become widely accepted that domains are the basic units of structure, function and evolution in proteins[1] and it is thought that proteins with complex functions evolve by fusion of genes coding for individual domains[2,3], the domains are not uniformly defined. Most commonly, domains are simply the compact and more or less loosely connected areas[1,4] apparent from a visual inspection of protein models; to avoid subjectivity and ambiguities inherent in visual inspection, certain computer algorithms for location of these 'structural' domains have

0028–0836/81/190085—03$01.00

recently been proposed[5–7]. An alternative interpretation is that domains are stable protein fragments found in biochemical experiments[8–17]. I regard them as 'globular fragments' which may refold autonomously and carry specific functions[1], and I propose here a method for location of these globular fragments based on surface area[18] measurements. Applied to several proteins the globular fragments found often coincide with structural domains or are contained within them. In particular the globular fragments found in globins correlate with the two structural domains proposed previously[4,19], and do not correlate with the three coding sequences separated by introns in the haemoglobin genes[20–22].

It has been shown that the surface area buried in a native fold is proportional to the gain in free energy of dehydration and hydrogen bond formation due to folding[23,24]. As surface area is proportional to the number of interatomic contacts[25], it may also serve as a crude estimate of van der Waals' interactions, which are very sensitive to errors in interatomic distances and therefore are usually averaged by counting the number of contacts. These findings are used here to define globular fragments.

A globular fragment may be defined as a segment of a protein chain which, when the rest of the protein is removed, buries a maximum of surface area within its fold. As the contribution of the free energy (estimated as 24 cal mol^{-1} Å^{-2} of buried surface area[24]) to its stability is maximal, such a fragment is more likely than others to be autonomously stable.

The surface area buried in stable globular proteins follows the law[24]

$$\begin{aligned} B_G &= 0.859\mathrm{MW} + 0.774\mathrm{MW}^2 \times 10^{-5} \\ &= 94.49L + 93.654L^2 \times 10^{-3} \qquad (1) \end{aligned}$$

which relates buried areas, B_G (in Å^2), to the molecular weight, MW, or the number of residues, L ($\mathrm{MW} \approx 110L$) in proteins of 50–320 residues. Non-globular structures have smaller buried areas (B) than those given by equation (1), the ratio B/B_G giving a measure of their globularity. Thus, we may calculate the buried surface areas of all chain fragments in a protein, removing the rest of the structure but assuming that the native conformation is retained, and analyse their B/B_G values. A maximum of B/B_G close to unity indicates the presence of a globular fragment.

Buried surface areas may be calculated from atomic coordinates[18]. The use of the simplified geometry[25,26] and the fast algorithm[7] allow such calculations to be made with efficiency equal to those for interatomic distances (see refs 5, 27). The buried areas thus determined are within a few per cent of those calculated using the all-atom algorithm[24]. Figure 1*a* shows, for the haemoglobin β chain, a significant difference in buried areas between globular and non-globular fragments of the same length. Fragment 19–68 has the buried area $B = 4{,}800$ Å^2 and $B/B_G = 0.97$, which indicates its globularity. The fragments of 50 residues beginning at residues 1–10 and 33–55 have $B = 3{,}500 \pm 250$ Å^2 and $B/B_G = 0.71 \pm 0.05$—these are non-globular. It has not been determined whether equation (1) is valid for $L < 50$. However, I found that conversion of the difference $B_G - B$, calculated for a number of α helices, to the equivalent free energy yields values only a few kcal mol^{-1} higher than the experimental stability of an α helix. Thus I used equation (1) for all chain lengths ($L \leq 320$), bearing in mind that it is less justified for $L < 50$. Figure 1*b* shows a typical B/B_G plot.

A summary of my results (Table 1) shows that globular fragments found here often coincide with structural domains or are contained within them. Ribonuclease is a notable exception: S protein (residues 21–124), found to be a globular fragment, was never listed previously among structural domains[4–7]. Table 1 also shows that my locations of globular fragments agree well with those determined experimentally. Note that location of the two independent globular fragments within the previously proposed entire discontinuous domain in hen lysozyme agrees with significant mutual displacements of these fragments on denaturation shown by X-ray crystallography[14]. A relatively low value of B/B_G for the C-terminal globular fragment of elastase is an artefact due to the unusually large number of small residues in this fragment—this means that the expression MW110L in equation (1) is invalid. The use of the dependence on MW in equation (1) yields $B/B_G = 0.98$ for this fragment, indicating

Table 1 Structural domains and globular fragments

Protein	Structural domains	Globular fragments: Calculated limits	B/B_G	Experiment
Pancreatic trypsin inhibitor	1–58*	1–55	1.10	
Immunoglobulin light chain	1–109†	4–109	0.96	Two globular fragments
	110–214†	110–209	0.96	similar to domains (refs 8–11)
Ribonuclease S	No continuous domain*	23–124	0.99	Globular fragment in 21–124 (refs 12–13)
Hen lysozyme	(1–38)+(101–129)*	8–29‡	1.05	8–29, 108–116 (ref. 14)
	39–87	39–98	0.97	42–99 (refs 14, 15)
		105–116‡	0.97	
T_4 lysozyme	1–74§	13–67	1.01	
	75–164§	78–154	1.05	
Papain	10–111*	17–100	1.11	Two independently melting
	113–207*	112–208	1.03	fragments (ref. 8)
Elastase	27–127*	27–120	1.07	
	128–230*	130–229	0.93‖	126–245 (ref. 16)
Thermolysin	1–157†	4–151	1.09	
	158–316†	203–316	1.06	206–316 (ref. 17)
Myoglobin	1–85*	21–70	1.03	
	86–153*	86–152	0.99	
Haemoglobin α chain	1–74*	20–63	0.91	
	75–141*	80–140	0.94	
Haemoglobin β chain	1–79*	19–68	0.97	
	80–146*	81–145	0.99	

Globular fragments are detected as maxima of B/B_G, where B is the buried surface area of a fragment and B_G is given by equation (1). If there are several maxima of B/B_G within the same chain region, only one fragment corresponding to the highest B/B_G value is listed.

$B = B_T - B_N$, where B_T is the total inaccessible area of all the spheres representing residues[25,26] in a given structure and B_N is the area of the spheres made inaccessible by their covalent neighbours along the chain in the same structure. Coordinates used in the calculations are from Brookhaven Protein Data Bank[28].

* Taken from ref. 4, haemoglobin chains are aligned to myoglobin. † Taken from ref. 35. ‡ Globular fragments much shorter than 50 residues (see text). § Located from stereodrawing in ref. 36. ‖ Low value due to an artefact; equals 0.98 when corrected (see text).

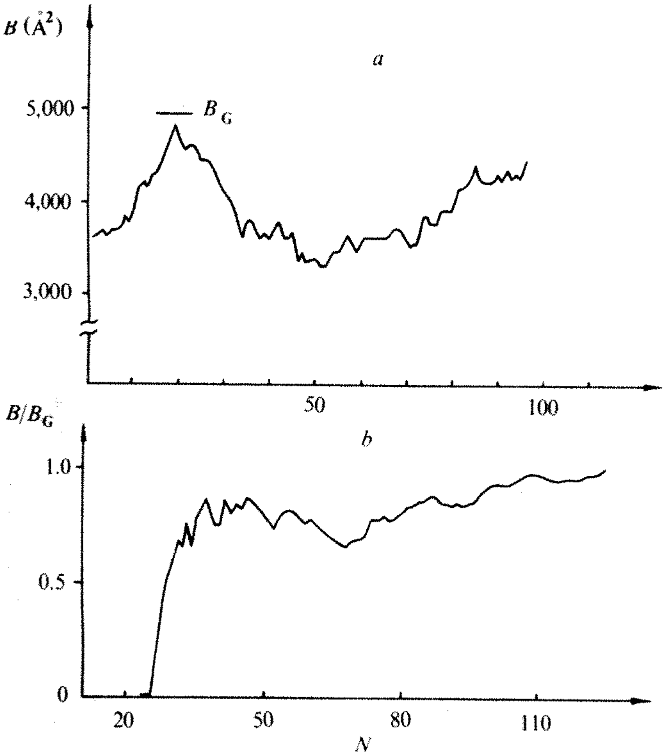


Fig. 1 *a*, The buried surface area B of each segment of 50 residues, beginning at residue N, is calculated for human β chain haemoglobin, assuming that the isolated segment retains its native conformation. The value B_G expected from equation (1) is almost reached by the segment beginning at $N = 19$. *b*, The ratios of the buried area values B of all segments of ribonuclease, beginning at residue 23 and ending at residue N, to the corresponding values B_G expected from equation (1). Two maxima with B/B_G values close to unity at $N = 110$ and $N = 124$ can be seen, the latter being slightly higher.

that this use may be preferable. B/B_G values for globin fragments calculated here using the simplified geometry corresponding to the time-averaged positions of the side chains[25,26] are higher than those calculated with the all-atom algorithm[18,24] (or the simplified geometry fitted to the actual positions of the side chains[7]). This effect of the change in side chain positions agrees with the necessity for small changes of the structure to gain more stability found in the model-building studies of globin fragments[29]. The lowest B/B_G value listed (which may be supposed to correspond to the lowest stability) is that for the N-terminal globular fragment of haemoglobin α chain. This reflects the lack of D helix in the α chain, which is the major difference between the latter and the β chain or myoglobin, and correlates with the drastically different conformation of the apo-α chain compared with the apo-β chain or the apomyoglobin[30].

It was recently proposed, by analogy with immunoglobulins, that translation products of the three continuous sequences or exons in haemoglobin genes correspond to domains[19–21]. Exon products for the β chain are fragments 1–30, 31–104 and 105–146. Table 1 shows that they correspond to neither structural domains[4,19] nor globular fragments. Exon products have low B/B_G values of 0.68, 0.85 and 0.81, indicating their nonglobularity. They do not seem to correspond to functional domains as the latter are conventionally viewed as being composed of few structural domains (or globular fragments)[1]. However, it was shown that fragment 31–104 binds cyanomethaemin but lacks any detectable structure even when bound to haem[31]. Alternatively, it has been proposed[32] that the function of each globular fragment in myoglobin is recognition of the haem. It may be speculated that residues 31–63, part of the N-terminal globular fragment, possess marginal stability sufficient to recognize the haem. We found that its B/B_G was 0.94 in the β chain, but this was not confirmed by the all-atom calculation, suggesting some rearrangement of the structure. This distorted structure may have low helicity undetectable by circular dichroism measurements. These results seem to indicate that at present the domain structure of globins cannot be unequivocally correlated with exons; therefore the problem needs further study.

Algorithms using surface area criteria have been successful in prediction of tertiary[33,34] and quaternary[25] structures of proteins and in location of structural domains (ref. 7 and Wodak and Janin, personal communication). I have shown here that they detect the presence within globular proteins of fragments that may be autonomously stable. Such fragments are likely candidates for intermediates in folding. The folding schemes based on surface area measurements can be derived[32] and compared with results obtained experimentally.

I thank B. H. Yudman for help in programming; M. Yu. Pavlov for control calculations using the all-atom algorithm; P. L. Privalov for useful discussion, and J. Janin and S. J. Wodak for useful discussions and for drawing my attention to the domain/exon hypothesis.

Received 9 July 1980; accepted 18 February 1981.

1. Schulz, G. E. & Schirmer, R. H. *Principles of Protein Structure* Ch. 5.3 (Springer, New York, 1979).
2. Edelman, G. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **63,** 78–85 (1969).
3. Rossmann, M. G., Moras, D. & Olsen, K. W. *Nature* **250,** 194–199 (1974).
4. Wetlaufer, D. B. *Proc. natn. Acad. Sci. U.S.A.* **70,** 697–701 (1973).
5. Crippen, G. M. *J. molec. Biol.* **126,** 315–332 (1978).
6. Rose, G. D. *J. molec. Biol.* **134,** 447–470 (1979).
7. Wodak, S. J. & Janin, J. *Proc. natn. Acad. Sci. U.S.A.* **77,** 1736–1740 (1980).
8. Privalov, P. L. *Adv. Protein Chem.* **33,** 167–241 (1979).
9. Rowe, E. S. & Tanford, C. *Biochemistry* **12,** 4822–4827 (1977).
10. Goto, Y., Azuma, T. & Hamaguchi, K. *J. Biochem.* **85,** 1427–1438 (1979).
11. Goto, Y. & Hamaguchi, K. *J. Biochem.* **86,** 1433–1441 (1979).
12. Labhardt, A. M. & Baldwin, R. *J. molec. Biol.* **135,** 245–254 (1979).
13. Chavez, L. G. & Scheraga, H. A. *Biochemistry* **19,** 996–1004 (1980).
14. Yonath, A., Podjarny, A., Honig, B., Sielecki, A. & Traub, W. *Biochemistry* **16,** 1418–1424 (1977).
15. Johnson, E. R., Anderson, W. L. & Wetlaufer, D. B. *J. biol. Chem.* **253,** 3408–3414 (1978).
16. Ghelis, C., Tempete-Gaillourdet, M. & Yon, J. M. *Biochem. biophys. Res. Commun.* **84,** 31–36 (1978).
17. Fontana, A., Vita, C. & Chaiken, I. M. *Hoppe-Seyler's Z. physiol. Chem.* **360,** 1006–1007 (1979).
18. Lee, B. K. & Richards, F. M. *J. molec. Biol.* **55,** 379–400 (1971).
19. Ptitsyn, O. B. & Rashin, A. A. *Biophys. Chem.* **3,** 1–20 (1975).
20. Gilbert, W. *Nature* **271,** 501 (1978).
21. Blake, C. C. F. *Nature* **277,** 598 (1979).
22. Eaton, W. A. *Nature* **284,** 183–185 (1980).
23. Chothia, C. *Nature* **248,** 338–339 (1974).
24. Chothia, C. *Nature* **254,** 304–308 (1975).
25. Wodak, S. J. & Janin, J. *J. molec. Biol.* **124,** 323–342 (1978).
26. Levitt, M. *J. molec. Biol.* **104,** 59–107 (1976).
27. Rossmann, M. G. & Liljas, A. *J. molec. Biol.* **85,** 177–181 (1974).
28. Bernstein, F. C. *et al. J. molec. Biol.* **112,** 535–542 (1977).
29. Rashin, A. A. in *Biomolecular Structure, Conformation, Function and Evolution* (ed. Srinivasan, R.) 133–149 (Pergamon, Oxford, 1978).
30. Yip, Y. K., Waks, M. & Beychok, S. *J. biol. Chem.* **247,** 7237–7244 (1972).
31. Craik, C. S., Buchman, S. R. & Beychok, S. *Proc. natn. Acad. Sci. U.S.A.* **77,** 1384–1388 (1980).
32. Rashin, A. A. *Studia Biophys.* **77,** 177–184 (1979).
33. Richmond, T. J. & Richards, F. M. *J. molec. Biol.* **119,** 537–555 (1978).
34. Cohen, F. E., Sternberg, M. J. E. & Taylor, W. *Nature* **285,** 378–382 (1980).
35. Liljas, A. & Rossmann, M. G. *A. Rev. Biochem.* **43,** 475–507 (1974).
36. Matthews, B. W. & Remington, S. J. *Proc. natn. Acad. Sci. U.S.A.* **71,** 4178–4182 (1974).

O_2 binding properties of the product of the central exon of β-globin gene

Charles S. Craik, Steven R. Buchman & Sherman Beychok*

Departments of Biological Sciences and Chemistry,
Columbia University, New York 10027, USA

The hypothesis that the exons of eukaryotic structural genes code for functional domains and that the partitioned arrangement of coding information may thus serve to mediate the rapid evolution of new and unique proteins from pre-existing exons[1–3] is also supported by our recent studies which demonstrate that the product of the central exon of the human β-globin gene is a complete functional domain capable of binding haem tightly and specifically[4,5]. Moreover, an analysis of the structure/function

---

* To whom correspondence should be addressed.

changes induced by mutations in different parts of the haemoglobin molecule suggests that each of the three exon-encoded segments is primarily associated with different functions of haemoglobin (for example, haem-binding, haem–haem interaction)[6]. We have now extended our studies to determine whether the central fragment is sufficient for maintenance of a stable complex of ferrous haem with molecular oxygen and, if not, what are the minimal requirements for the expression of this activity. The results of our reconstitution experiments indicate that the isolated central exon peptide is unable to maintain a ferrous haem–dioxygen complex unless the side exon products and the complementary haem-containing subunit are present. A conformational change which accompanies the noncovalent association of fragments may account for the restoration of reversible oxygen binding.

Haemoglobin and myoglobin in their ferrous states reversibly bind oxygen—this makes them unique among haem-binding proteins and different from most ferrous haem compounds, which are rapidly oxidized in the presence of molecular oxygen. They are the only molecules known, other than the biomimetic model systems (for example, haem–polystyrene matrices[7] and 'picket fence' porphyrins[8]) that allow O_2 and Fe(II) haem to combine and dissociate freely without concomitant oxidation of the iron. The stability of the oxygenated derivative seems to be achieved in substantial part by accommodation of the haem group in a hydrophobic pocket into which the side chains from the surrounding helices project. These side chains provide the fifth ligand (His F8), modify the binding of the sixth ligand (O_2, CO, CN^-, etc.) and lock the haem into a specific spatial arrangement with respect to the rest of the protein[9,10].

The distinctive spectral characteristics of various ligand derivatives of haemoglobin have been well documented and provide a sensitive means of evaluating the environment and ligation state of the prosthetic group in the protein[11]. The method used to investigate the oxygen binding properties of the globin central fragment (see Fig. 1) was based on spectral analysis; its application to the isolated intact globins, after reconstitution with haem, was used as a control. Figure 2*a* shows spectra resulting from a typical oxygenation experiment in which the intact β-globin chain (β^0) was combined with haemin dicyanide to form the cyanmet subunit derivative, reduced with sodium dithionite and subsequently oxygenated. Spectra for α-subunits of haemoglobin, native or reconstituted, are almost identical to those presented for β, differing only slightly in the absolute value of the extinction coefficients and having a 1–2 nm difference in position of the absorbance maxima.

Table 1 Far UV circular dichroism and secondary structure characteristics

Protein or peptide	$[\theta]_{222\,\mathrm{nm}}$ (deg cm^2 per dmol)	Calculated % α helix*	Calculated no. of helical residues†
β^h	21,600	73	106
β^0+h	21,600‡	73	106
$\beta^0_{digest}+\alpha^h+h$	18,200		
	(14,800)§	52§	76
$\beta^0_{rec}+\alpha^h+h$	17,500		
	(13,400)§	47§	69
β^0	10,600	39	57
$\beta^0_{digest}+h$	10,500	38	55
$\beta^0_{rec}+h$	9,500	35	51
β^0_{digest}	7,300	28	41
β^0_{rec}	7,500	29	42
$\beta^0_{cf}+\alpha^h+h$	13,700		
	(5,800)§	23§	17
β^0_{cf}	5,800	23	17
α^h	22,000	78	110
α^0+h	21,500‡	73	103
$\alpha^0_{digest}+\beta^h+h$	18,900		
	(16,200)§	57§	79
$\alpha^0_{digest}+h$	14,000	49	69
α^0	11,000	40	56
α^0_{digest}	7,200	28	39

α^0 or β^0 indicates that the chain has no haem attached; α^h or β^h indicates that the chain has bound haem. The subscript 'digest' refers to unfractionated clostripain-treated globin; the subscript 'rec' refers to a fully reconstituted sample prepared by mixing together equimolar concentrations of the various arginyl peptides, and the subscript 'cf' denotes the isolated central fragment of β-globin. $[\theta]_{222}$ values are means of at least three determinations, except for β^0_{rec} which is the mean of two determinations. The estimated uncertainty in $[\theta]_{222}$ is ±5%. In the cases of the β^0_{rec} samples, the second digit in the calculated α helix may not be significant due to uncertainty in concentration determination.

* Calculated according to Chen *et al.*[21].

† Calculated by multiplying the fraction of helical residues of the third column by 146 and 141 for α- and β-globins, respectively, and 74 for β^0_{cf}.

‡ Reconstitution of α- and β-globins with haemin dicyanide to give the same $[\theta]_{222}$ requires dialysis into 0.1 M potassium phosphate buffer, 1 mM EDTA, *p*H 7.4.

§ These values refer to the β components after subtracting the contribution made by α^h.

To follow oxygen binding to the haem-binding domain, by itself and in combination with other components of haemoglobin, the α- or β-globin chain was first proteolytically cleaved with the arginine-specific protease into the fragments corresponding to the exon-encoded regions of the protein (Fig. 1). The visible and Soret spectra of these samples after their reconstitution with haemin dicyanide were then obtained along with their reduced and oxygenated derivatives. Comparison of Fig. 2*b* with the intact control shown in Fig. 2*a* reveals almost identical spectra for haem-reconstituted β-globin and β-globin digest. However, whereas reduction with dithionite produced a deoxyhaemoglobin spectrum for the reconstituted β-globin, a haemochromogen spectrum (hexacoordinate ferrous haem iron axially bound to two histidines) was obtained with the reconstituted digest and with the isolated central fragment. Subsequent removal of the dithionite produced an oxyhaemoglobin spectrum for the undigested samples; in contrast, the spectrum of the digested sample or the central fragment was that of a haemichromogen (hexacoordinate ferric haem with two histidines serving as the axial ligands), diagnostic of a haem complex that does not reversibly bind oxygen. Similar results were found for α-globin.

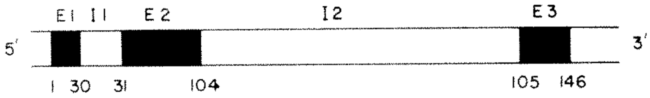


Fig. 1 Diagrammatic representation of the β-globin intron/exon structure. The black boxes E1, E2 and E3 represent expressed DNA sequences (exons); the white boxes I1 and I2 represent intervening sequences (introns)[16]. The open-boxes at the 5′ and 3′ ends are untranslated regions of indeterminate length. The numbers below the exons denote amino acid sequence positions of the β-globin. E2 represents the central exon that codes for the haem-binding polypeptide consisting of amino acids 31–104. The 5′→3′ orientation is indicated. A similar arrangement exists for the α-globin gene[17,18]. The small intron (I1) for α-globin occurs between codons 31 and 32 whereas the large intron (I2) interrupts the coding sequence at 99 and 100. When the amino acid sequences of α- and β-globin are aligned to maximize functional homology[19], their intron/exon junctions coincide; α-globin residue 31 is an arginine but residue 99 is a lysine, thus limited clostripain digestion yields only two fragments of residues 1–31 and 32–141. Cleavage products were separated by a modification of the procedure used elsewhere[4] and involved gel permeation high performance liquid chromatography in 6 M urea, 0.2M formic acid, using a recycle purification mode[20]. This procedure is much more rapid and gives higher purity than molecular sieve chromatography in 9% formic acid.

Comparable experiments were carried out on a fully reconstituted sample of β-globin digest (β^0_{rec}), prepared by mixing equimolar concentrations of the various arginyl peptides then combining the mixture with haemin dicyanide. Although the mixture showed a substantial Soret extinction coefficient, haemochromogen and eventual haemichromogen spectra resulted from dithionite addition and its subsequent removal. However, when the complementary haem-containing subunit, α^h, was combined with β^0_{rec} or the β^0_{digest} before haem addition, the typical oxy spectrum presented in Fig. 2*c* resulted. By subtracting the contribution to this oxy spectrum due to the native α^h, it was established that 86% of the β^0_{digest} produced an oxy spectrum, with the other 14% haemichromogen. The β^0_{rec}

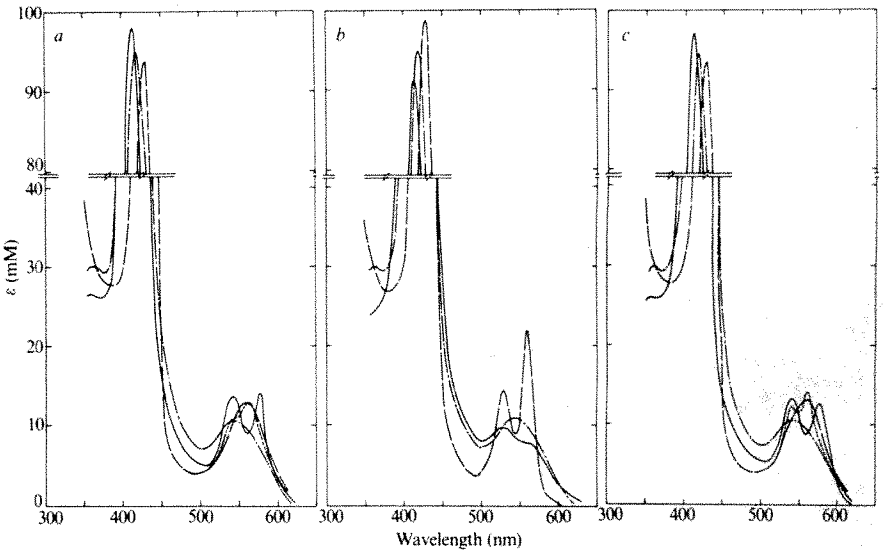


Fig. 2 Spectral properties of cyanmet-, deoxy- and oxy-derivatives of undigested and digested globins. *a*, Cyanmet-spectrum (— — —) of an equimolar mixture of *β*-globin and haemin dicyanide; transient ferrocyanide spectrum (---------) immediately after sodium dithionite addition and the eventual deoxy-spectrum (— — —); oxy-spectrum (———) after dithionite removal. The transient ferrocyanide spectrum and the deoxy-spectrum are coincident below 500 nm. Solutions of the ferricyanide derivatives showed an absorption spectrum in the visible with a broad band centred at 540 nm and a barely discernible shoulder at 560 nm. The Soret band lies near 420 nm, and has an extinction coefficient about 10-fold greater than that at 540 nm. Ferricyanide derivatives of haemoglobin react directly with dithionite, leading to reduction of the haem iron with ligand still bound, the transient ferrocyanide derivatives then dissociating into cyanide and deoxygenated ferrous haem-protein. The latter transition from ferrocyanide to deoxyferrohaem could be monitored for the separated *α*- and *β*-subunits, as the ferrocyanide spectrum, with a maximum at 562 nm and a smaller peak at 541 nm, only gradually coalesced to the deoxyferrohaem spectrum, characterized by a single, broad asymmetric band in the visible region with a maximum at 556 nm and a Soret band in the near UV at ~430 nm. After removal of dithionite by dilution, dialysis or gel filtration, and exposure of the haem–protein solution to oxygen, the characteristic oxyferrohaem spectrum appears. The two visible absorption bands are quite distinct from each other, with one peaking at 577 nm and a broader one of slightly lower intensity peaking at 541 nm. The Soret band is at 415 nm. *b*, Cyanmet-spectrum (— — —) of an equimolar mixture of 60-min clostripain digested *β*-globin and haemin dicyanide; haemochromogen spectrum (— — —) due to dithionite addition and a haemichromogen spectrum (———) after dithionite removal. For *α*-globin, this spectrum also contains a 9.3% oxy-component presumably due to the small amount of undigested globin present. *c*, Cyanmet-spectrum (— — —) of an equimolar mixture of 60-min clostripain digested *β*-globin, the complementary haem-containing *α*-subunit and haemin dicyanide; transient ferrocyanide spectrum (---------) immediately after dithionite addition and the eventual deoxy-spectrum (— — —); oxy-spectrum (———) after dithionite removal. The transient ferrocyanide spectrum and the deoxy-spectrum are coincident below 500 nm. Measurements were made at 4 °C in 20 mM potassium phosphate, 1 mM EDTA buffer, *p*H 5.7 or 0.1 M potassium phosphate buffer, 1 mM EDTA buffer, *p*H 7.4. Although not essential for α^0 reconstituted with haem, 0.1 M potassium phosphate *p*H 7.4 buffer was required to develop reversible oxygen binding activity for the haem-reconstituted *β*-subunit. To reduce undesirable side effects, dithionite reductions were achieved with a 10-fold excess of dithionite in a nitrogen atmosphere. Dithionite removal was by gel filtration through a Bio-Gel P-2, dialysis or dilution.

sample yielded a lower value of 74%. This discrepancy was probably due to irreversible denaturation of the fragment produced by the separation procedures.

No appreciable oxy spectrum was observed when the experiment was repeated with the central fragment alone mixed with α^h. It is therefore clear that the reversible O_2 binding abilities of the central exon haem-binding fragment, incapable of reversibly binding O_2 on the time scale (min) and in the conditions of the above experiments, can be restored by the presence of the side exon fragments and the complementary haem-containing subunit.

To determine whether reversible O_2 binding is correlated with secondary structure, the far UV circular dichroism spectra of the haem-free globins, the fractionated, unfractionated and reconstituted digestion products, the central fragment of *β*-globin and their haem-reconstituted mixtures were determined; the results are presented in Table 1. The helix contents of both *α*- and *β*-globins, which are half those of the haem-containing chains, or lower depending on conditions, were reduced further by digestion with clostripain to the extent of an additional 10% loss in helicity. Both α^0 and β^0 digests, as well as β^0_{rec}, showed some recovery of helix content on recombination with haemin dicyanide: ~20% (or 30 residues) refolding in the case of α^0_{digest} and ~10% in the case of β^0_{digest} and β^0_{rec}. The addition of the complementary haem-containing subunit α^h to β^0_{digest} or β^0_{rec}, and β^h to α^0_{digest} induced some refolding, ~10% in the case of α^0_{digest} and 15% in the case of β^0_{digest}. These changes are smaller in magnitude than those seen when α^h is added to *β*-globin or β^h is added to *α*-globin, perhaps reflecting the limitations imposed on refolding by the interruption of covalent bonding in the linear sequence of residues of the native helices. Neither addition of haemin dicyanide or α^h affected the circular dichroism of the isolated β^0 central fragment.

The loss of structure of the globins on clostripain cleavage, as seen from the circular dichroism measurements, seems to be sufficient for coordination of the distal histidine with the haem iron, thus preventing oxygen binding and resulting in haemichromogen formation. Once the ability to restore the native three-dimensional structure of the haem-binding cavity has been lost, the maintenance of a stable ferrous complex in the presence of oxygen is prevented. This is clearly the case for haemoglobin or myoglobin where the stability of the oxygenated derivative is lost on denaturation, linking the integrity of the conformation of the protein to its O_2-binding capabilities.

The possibility that the structural features defining the haem pocket are induced by haem-binding to the central fragment and can be made more precise when the side fragments are present, even if not covalently joined, has been suggested from Soret extinction coefficient measurements of haem-reconstituted β^0, β^0_{digest}, β^0_{rec} and the central fragment. However, this was not sufficient to restore secondary structure or reversible oxygen binding to the haem-binding fragment. Only by introducing equivalent amounts of the complementary haem-containing subunit before reconstitution with haem could an oxyhaemoglobin spectrum result. This was accompanied by a 14% increase in helix content (20 residues) which suggest a restructuring of the haem pocket to allow reversible O_2 binding. The phenomenon is reminiscent of the refolding observed in complementation studies on ribonuclease[12] and staphylococcal nuclease[13].

These results, taken together with our earlier studies, demonstrate that the exon-encoded central fragment has the structural potential of providing a tight and specific binding site for haem, and that the fit is sharpened by the addition of the N- and C-terminal fragments, but that formation of a stable dioxygen complex requires, in the conditions of our experiments, the presence of the side fragments and the complementary subunit.

We are investigating, by stop-flow methods, whether transient reversible oxygen binding in the absence of the complementary subunit occurs, to elucidate the mechanism by which the distal histidine becomes irreversibly bound to the iron. Obviously, the requirement of a complementary subunit cannot be general, as shown by the reversible oxygen binding of the *α* subunit and myoglobin, both monomeric. In this case, the added subunit served to overcome the tendency for the distal histidine to be bound to haem in the deoxy state. However, it is precisely a difference in bonding tendency of the distal histidine that distinguishes the globins from the *b*-type cytochromes[11]. The central exon product, isolated or in noncovalent association with the side peptide fragments, accordingly has spectral characteristics reminiscent of these cytochromes, and undergoes a similar reduction–reoxidation cycle with dithionite and molecular oxygen.

The finding of a haem-binding (or, more generally, porphyrin-binding) domain encoded by an exon that can combine with other gene elements to generate a multiplicity of haem-binding proteins of diverse function, would be very fundamental. Thus we have begun to define the essential features of a haem-binding site and have shown similarity in the nature and spatial distribution of haem contact residues in haemoglobin chains and several *b*-type cytochromes[5,14,15]. We will provide details of these similarities elsewhere and suggest that the evolution of the globin genes from earlier cytochrome-like ones involved combination of an exon coding for a haem-binding domain with others not derived from cytochrome genes.

This study was supported in part by NSF grant PCM 80-05614. We thank Drs Alex Tzagoloff and Jonathan Greer for critical reading of the manuscript and constructive advice, and Drs Jeremy Luchins and Yvonne Leutzinger for their interest and support. The authors also thank Ms Jill Macoska for preparation of the text, table and figures.

Received 15 January; accepted 2 March 1981.

1. Gilbert, W. *Nature* **271,** 501 (1978).
2. Blake, C. C. F. *Nature* **277,** 598 (1979).
3. Gilbert, W. in *Eucaryotic Gene Regulation* (eds Axel, R., Maniatis, T. & Fox, C. F.) 1–12 (Academic, New York, 1979).
4. Craik, C. S., Buchman, S. R. & Beychok, S. *Proc. natn. Acad. Sci. U.S.A.* **77,** 1384–1388 (1980).
5. Craik, C. S. & Beychok, S. in *Interaction Between Iron and Proteins in Oxygen and Electron Transport* (ed. Chien Ho) (Elsevier North-Holland, New York, in the press).
6. Eaton, W. *Nature* **284,** 183–185 (1980).
7. Wang, J. H. *Acc. Chem. Res.* **3,** 90–97 (1970).
8. Collman, J. P. *Acc. Chem. Res.* **10,** 265–272 (1977).
9. Perutz, M. F. *A. Rev. Biochem.* **48,** 327–386 (1979).
10. Perutz, M. F. *Scient. Am.* **239,** 92–125 (1978).
11. Lemberg, R. & Legge, J. M. *Hematin Compounds and Bile Pigments* (Interscience, New York, 1949).
12. Brown, J. E. & Klee, W. A. *Biochemistry* **10,** 470–476 (1971).
13. Taniuchi, H., Anfinsen, C. B. & Sodja, A. *Proc. natn. Acad. Sci. U.S.A.* **58,** 1235–1242 (1967).
14. Leutzinger, Y. & Beychok, S. *Proc. natn. Acad. Sci. U.S.A.* **78,** 780–784 (1981).
15. Nobrega, F. G. & Tzagoloff, A. *J. biol. Chem.* **255,** 9828–9837 (1980).
16. Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G. & Maniatis, T. *Cell* **15,** 1157–1174 (1978).
17. Nishioka, Y. & Leder, P. *Cell* **18,** 875–882 (1979).
18. Lauer, J., Shen, C.-K. J. & Maniatis, T. *Cell* **20,** 119–130 (1980).
19. Dickerson, R. E. & Geis, I. *The Structure and Action of Proteins*, 52 (W. A. Benjamin, California, 1969).
20. Craik, C. S. thesis, Columbia Univ. (1980).
21. Chen, Y.-H., Yang, J. T. & Chau, K. H. *Biochemistry* **13,** 3350–3359 (1974).

Correlation of DNA exonic regions with protein structural units in haemoglobin

Mitiko Gō

Department of Biology, Faculty of Science, Kyushu University 33, Fukuoka 812, Japan

The discovery of intervening sequences (introns) in DNA led Gilbert and Tonegawa[1,2] to suggest that a new protein could have been produced by bringing together certain segments of pre-existing ones. However, Blake[3] argued that if DNA was so organized that coding sequences (exons) correspond to structural as well as functional units of proteins, then combinations would be much more likely to yield a stable globular conformation through being 'sums of parts'. In immunoglobulin heavy chain, four separate exons encode four different units[4], all with distinct functions and three of which have clear domain structures. However, in haemoglobin, which has no obvious domain structure, no clear conformational characteristics have so far been recognized for the segments encoded by exons. From a close inspection of their conformations by drawing various stereodiagrams and the C^{α}–C^{α} distance map, I now propose a conformational characterization of the segments as structural units.

In the mouse α- and β-globin genes, the three coding sequences correspond to amino acid residues 1–31, 32–99 and 100–141, and 1–30, 31–104 and 105–146, respectively[5–8]. These two subunits of haemoglobin are structural homologues. The central coding sequence has been shown to correspond to the function that interacts with and binds haem[9,10]. Do exonic regions of DNA correspond to structural units of protein conformation in haemoglobin? It is well known that the globin monomers do not show domain or supersecondary structures[11] such as those observed in immunoglobulin or dehydrogenases. A domain can be regarded as a part of a protein which forms a compact, continuous globular region separated from the rest of the molecule[12,13]. The joints at residues 30–31 (31–32 in the β chain) and 104–105 (99–100 in the β chain) between adjacent segments encoded by different exons are located in the middle of B and G helices, respectively. Thus the exonic regions do not correspond to simple secondary structures. Therefore, the conformations of the segments are not characterized as domains, supersecondary or secondary structures.

The folding structure of the whole β chain and those of its three segments encoded by the exons are shown in Fig. 1*a–d*. The segment coded by the right exon consists of two helices connected by several residues forming a reverse turn (Fig. 1*b*). This segment is in slight contact with the haem group. The segment encoded by the central exon compactly encloses the haem group; 14 out of 17 side chains in contact with the haem are part of this segment (Fig. 1*c*). This correspondence of haem binding to the central coding region has been reported elsewhere[9,10]. The segment encoded by the left exon is a helical arc and is less compact than the other segments (Fig. 1*d*)—this segment is not in contact with the haem group. The mutual arrangements of the segments in Fig. 1*a* are such that they lie, and are thus exposed on, the surface of the molecule. They may thus be regarded as a surface layer of the molecule. There is no core in haemoglobin subunits, that is, the subunits consist solely of the surface layer, which is due to the small size of the subunits. Note also that in Fig. 1*a* the joints between adjacent segments occur relatively close to the centre of the molecule. In fact the residue accessibility to water molecules is extremely small at these joints[14].

Protein folding can be presented on a two-dimensional plane by the so-called C^{α}–C^{α} distance map[15] first used by Rossmann and Liljas[13] to locate folding domains. By careful inspection of various patterns on the distance maps, Kuntz[16] showed that the square and trapezoidal patterns represented the simplest and smallest units of the tertiary structure consisting of the three segments that are situated close together in space. Folding of the haemoglobin β chain is shown on this distance map in Fig. 2. The pairs of residues separated by >27 Å are shown by dark shading: these regions are grouped into seven major continuous areas. In Fig. 2 solid lines are drawn at the joints between the polypeptide segments encoded by different exons, that is, between residues 30 and 31 and between 104 and 105. It is remarkable that these lines scarcely cross the dark regions on the map.

This fact together with the characteristics observed in the stereodiagrams leads to a clear characterization of the conformations of the segments encoded by the exons. To describe this characterization, consider the following problem. Given the three-dimensional structure of each subunit of haemoglobin as shown in Fig. 1*a*, when cut into three or four segments consisting of 30–50 residues, in such a way that they are as little extended as possible, it is obvious that all conformations of such segments cannot be very compact. One or two must be rather extended. A possible algorithm for solving this problem in such small one-layer proteins as haemoglobin subunits uses pairs of residues

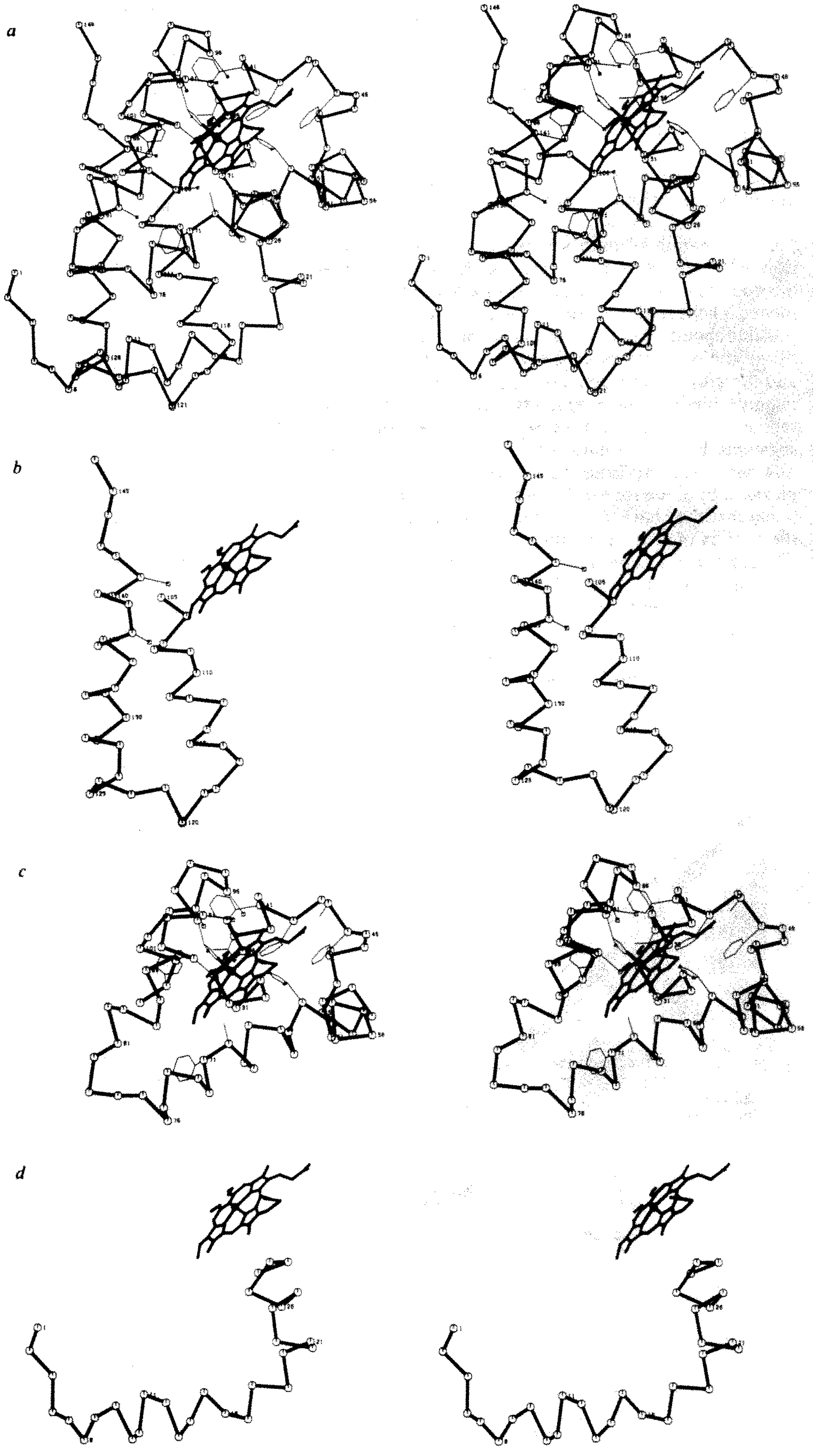


Fig. 1 Stereodiagrams of the whole β chain (*a*) and of the segments corresponding to the right (*b*), central (*c*) and left exons (*d*). The α-carbon backbone is represented. The haem group is shown with each segment to give the relative three-dimensional position. The side chains in contact with the haem group[20] (located within ~4 Å) are drawn symbolically.

which are very far apart in space and cut the polypeptide chain in such a way that pairs of residues that are far apart are not included within each individual segment. This is exactly the situation seen in Fig. 2 if we introduce one more cut somewhere between residues 66 and 71. A set of four segments, F1–4, presents a solution to our problem. F1, F2+F3 and F4 are encoded by exons. From inspection of Fig. 1*a* the central segment itself is seen to be very little extended. Thus, conformations of segments encoded by exons are characterized by the fact that they are the least extended of all possible choices of cuts in the polypeptide chain consisting of three or four segments of 30–50 residues. Pairs of residues which are far apart usually occur at two opposite sides of a globular structure. Therefore, cuts which are introduced to separate them tend to occur close to the centre of the molecule (Fig. 1*a*).

The structural units can be defined by the present algorithm for proteins whose X-ray data are available. Lysozyme is one of these proteins—its gene sequence has been reported by Jung *et al.*[17]. It will be reported elsewhere that three out of four junctions between adjacent structural units of lysozyme determined by the present algorithm correspond to the exon–intron junctions.

Conformation of each segment would not be stable when isolated[18] but the ability to assume 'the native format', when incorporated into a globular structure together with other segments, is kept at a maximum by 'being least extended' among all possible choices. In a small one-layer protein molecule, one side of the conformation of each segment is exposed to the surface and therefore should be mainly polar; the other side should be mainly nonpolar. The probability of forming a new protein from segments of old proteins is expected to be fairly high, when segments have the characters described above. The conformational characteristics now recognized for the segments encoded by the exons in haemoglobin subunits make it more likely that they have a structural, as well as a functional role in the process of protein evolution.

From the structural point of view it is more natural that the structural units F2 and F3 are encoded by separate exons. However, no intervening sequence has been found in the DNA at a position corresponding to the F2–F3 junction. One possible explanation for this absence is that selection pressure eliminated the intervening sequence that was present in an ancestral protein[19]. This is likely because the structural units F2 and F3, when put together, fix the haem group in its functional position, and because unnecessary exon shuffling promoted by presence of the intron would be harmful. In connection with this, note that an interesting feature of the amino acid sequences is observed when rather arbitrary cuts between F2 and F3 are introduced between residues 61 and 62 in the α chain and between 66 and 67 in the β chain (Fig. 3). The last residues of F1 in both chains, and of F3 in the α chain are common in many species—these are arginine or lysine. Arginine is the main residue at the last position of F3 in β chain—lysine is less common. The last residue of the F2 in both chains is commonly lysine. Furthermore, the first two residues of F2 and F4 are hydrophobic, which is again the case for the first two residues of F3. Although it is not known why lysine and arginine, followed by hydrophobic residues, are found at the joints between the segments encoded by exons, the common character of the amino acid residues at the three joints of the segments supports the possibility of the presence in, and subsequent elimination from the ancestral globin, of an intron between DNA sequences encoding the structural units F2 and F3.

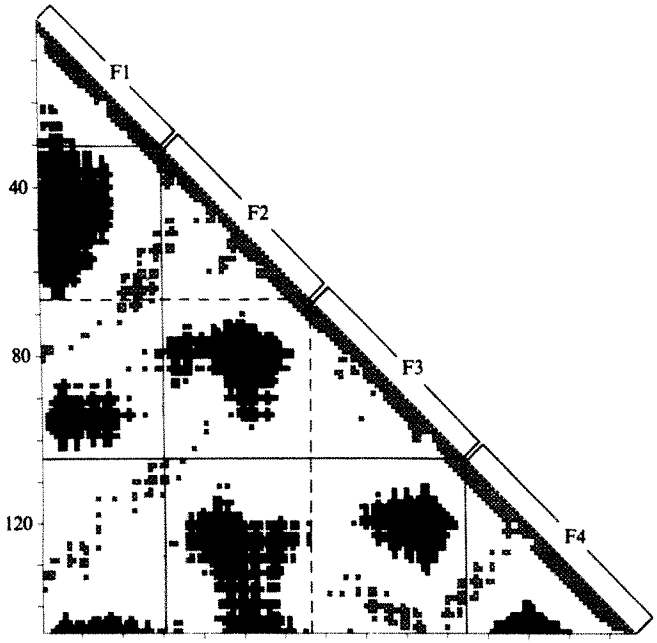


Fig. 2 Distance map of haemoglobin β chain. The distances between the *i*th and *j*th α-carbon atoms are plotted in a triangular space, in which both ordinate and abscissa are the residue sequence number. The grey region represents the pairs of C^{α} atoms within 9 Å, the open area the pairs whose distances are 9–27 Å, and the dark region the pairs separated by >27 Å. α-Helices are represented by a series of interactions along the diagonal. The structural units F1–F4 are defined by considering the exonic regions, this map and the amino acid sequences near the joints of the units. The units F1, F2+F3 and F4 are encoded by the three exons, respectively. The atomic X-ray data used are of horse deoxyhaemoglobin (provided by the Protein Data Bank at Brookhaven National Laboratory).

Structural unit	-F1	F2			F3		F4	
α chain								
Residue no.	---31	32	33---	---61	62	63---	---99	100	101--
Common	Arg			Lys			Lys		Leu
Major		Met	Phe		Val	Ala		Leu	
β chain									
Residue no.	---30	31	32---	---66	67	68--	--104	105	106--
Common	Arg	Leu	Leu	Lys	Val			Leu	Leu
Major						Leu	Arg		

Fig. 3 The amino acid residues that are common and major in various species[21] at the junctions of the structural units in haemoglobin α and β chains.

I thank Mr H. Mizuno for the use of the computer programs for the distance map and stereodiagram, Professor H. Matsuda for his encouragement during the work, Dr N. Gō for critical reading of the manuscript and various discussions, and Professor G. Schütz for providing the lysozyme exon–intron structures before publication. Part of this work was supported by grants-in-aid from the Ministry of Education, Japan.

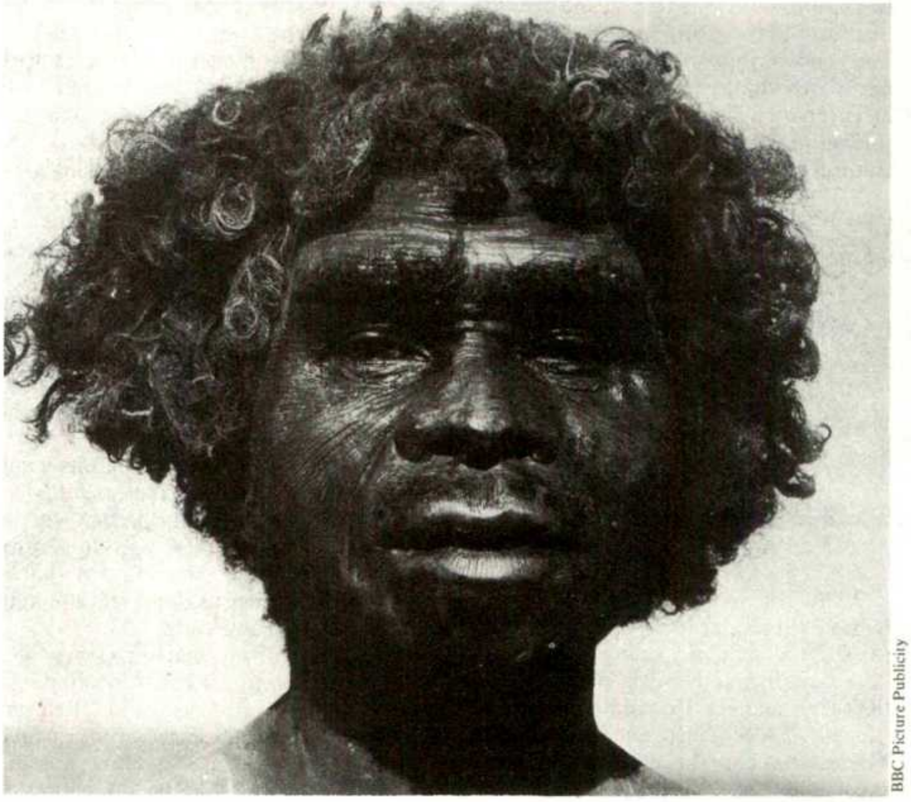
Received 25 September 1980; accepted 16 February 1981.

1. Gilbert, W. *Nature* **271,** 501 (1978).
2. Tonegawa, S., Maxam, A. M., Tizard, R., Bernard, O. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **75,** 1485–1489 (1978).
3. Blake, C. C. F. *Nature* **273,** 267 (1978).
4. Sakano, H. *et al. Nature* **277,** 627–633 (1979).
5. Leder, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **75,** 6187–6191 (1978).
6. Nishioka, Y. & Leder, P. *Cell* **18,** 875–882 (1979).
7. Konkel, D. A., Tilghman, S. M. & Leder, P. *Cell* **15,** 1125–1132 (1978).
8. Konkel, D. A., Maizel, J. V. Jr & Leder, P. *Cell* **18,** 865–873 (1979).
9. Craik, C. S., Buchman, S. R. & Beychok, S. *Proc. natn. Acad. Sci. U.S.A.* **77,** 1384–1388 (1980).
10. Eaton, W. A. *Nature* **284,** 183–185 (1980).
11. Rao, S. T. & Rossmann, M. G. *J. molec. Biol.* **76,** 241–246 (1973).
12. Wetlaufer, D. B. *Proc. natn. Acad. Sci. U.S.A.* **70,** 697–701 (1973).
13. Rossmann, M. G. & Liljas, A. *J. molec. Biol.* **85,** 177–181 (1974).
14. Gō, M. & Miyazawa, S. *Int. J. Peptide Protein Res.* **15,** 211–224 (1980).
15. Nishikawa, K., Ooi, T., Isogai, Y. & Saito, N. *J. phys. Soc. Japan* **32,** 1331–1337 (1974).
16. Kuntz, I. D. *J. Am. chem. Soc.* **97,** 4362–4366 (1975).
17. Jung, A., Sippel, A. E., Grez, M. & Schütz, G. *Proc. natn. Acad. Sci. U.S.A.* **77,** 5759–5763 (1980).
18. Anfinsen, C. B. & Scheraga, H. A. *Adv. Protein Chem.* **29,** 205–300 (1975).
19. Doolittle, W. F. *Nature* **272,** 581–582 (1978).
20. Perutz, M. F., Muirhead, H., Cox, J. M. & Goaman, L. C. G. *Nature* **219,** 131–139 (1968).
21. Dayhoff, M. O., Hunt, L. T., McLaughlin, P. J. & Barker, W. C. *Atlas of Protein Sequence and Structure* Vol. 5 (ed. Dayhoff, M. O.) D370–D371 (National Biomedical Foundation, Maryland, 1972).

BOOK REVIEWS

Not by bones alone: "The Making of Mankind"

Barry Cox


BBC Picture Publicity

Face from the past — an actor wearing the *Homo erectus* mask.

The Making of Mankind. By Richard E. Leakey. Pp. 256. ISBN 0-7181-1931-2. (Michael Joseph: 1981.) £9.95. Dutton will publish the book in the USA in September, price $25.75. The first programme in the seven-part television series was shown on BBC 2 on 4th May and the remainder will appear on successive Mondays at 9.50 pm.

IN the now common double-barrelled "Let 'em have it in the eyes and ears together" approach, the publication of Richard Leakey's book has accompanied the beginning of the seven-part BBC TV series of the same title — *The Making of Mankind*. After seeing a preview of the first and third programmes, reading the book allows one to make a reasonable guess at the nature of the complete series. I think, on the whole, that Richard Leakey deserves a high mark for balance, content and style.

The linkman of a series of programmes of this kind has an immense opportunity. Taken wherever there is something worth filming, with researchers to help in identifying and gathering relevant material, he can usefully indulge that desire to read and think beyond the boundaries of his own specialist research that lies in many scientists. Leakey has grasped this opportunity with both hands; he quotes Alan Walker's words: "If all we were to learn from prehistory was a list of species in time and space, then we should give up". *The Making of Mankind* is not, therefore, merely an anatomical/stratigraphical account of how one type of leg-bone evolved into another, but is an attempt to weave together all the different threads of evidence that may help us to understand the pathway that led to man today. This approach also has the advantage for the television viewer that he does not have the uneasy feeling that he must, from week to week, remember what went before or risk misunderstanding later programmes.

The beginning of such a series is always the most difficult to design, since one has to capture the interest of the audience and at the same time give an idea of the scope of the whole and provide some of the inevitably rather unglamorous chunks of basic background. Leakey starts with attractive film of exploration and discovery in East Africa, and then back-tracks to explain the processes of evolution and fossilization, and the geological time-scale. I think it was unnecessary to start with the stock sequence (photographs of volcanic activity — origin of the world — origin of life — metazoa — life on land etc.), and I also cannot see that the current debate over graduated evolution versus punctuated equilibria throws any light on human evolution. But the animation–explanation of the process of fossilization was extremely clear, and well integrated with a real-life simulation of rain-water revealing a fossil skull. The programme ends with an account of the earliest hominoid fossils, the forest dryopithecines and the woodland ramapithecines, and suggests how climatic change and consequent environmental change may have moulded this stage of human evolution. On the whole, I felt that this programme negotiated the initial problems very well.

The second programme looks at the African discoveries of fossils of *Australopithecus* and early types of *Homo*, and at the evidence that at least two species of hominid were apparently able to co-exist. It therefore includes the problematical Hadar fossils and the differing interpretations of Leakey and Johansen over their nature: do they all belong to one species of *Australopithecus*, or do some belong to an early species of *Homo*? Though Leakey has been accused of giving a one-sided view of the debate, that is certainly not true of this section of the book, while Don Johansen himself presents his own case in the TV programme, so I feel that the accusation is quite unfair — and the more regrettable for encouraging the layman's view that scientists cannot hold differing theories without becoming personally involved.

It is the third programme that Leakey starts to widen the scope of the series, emphasizing the new experimental approach to the problems of understanding the making and use of flint tools and the social significance of the camp-sites of early man. Here, too, he introduces his main message: that the fundamental social characteristic of man is that he is a *sharer*, his camp-sites being the locations to which food was brought to be shared. This theory, explained by Glynn Isaac, suggests that it was the contribution of both sexes to the food-sharing group (the males scavenging for meat while the females collected plant foods) that provided the basis for the eventual evolutionary superiority of *Homo* over *Australopithecus*. This way of life is presented dramatically by the use of actors made to resemble *Homo erectus* by wearing startlingly life-like masks, and by film of the modern !Kung bushmen who today follow the hunter–gatherer way of life into which this food-sharing habit may have evolved.

In later chapters and programmes, Leakey outlines recent work on the ability of apes to use signs and, by implication, to classify objects and concepts, and on the brain casts of fossil man. (The tendency to suggest that a larger brain *allows* more complex behaviour surely puts the morphological cart before the functional horse?). Neanderthal man leads on to Ice Age art; film cameras were allowed inside Lascaux for the first time in 20 years, and we are promised "staggering and beautiful images of horse, deer, bison and stags". It also seems as though horses and cattle may have been domesticated far earlier than had been thought. Leakey ends with a consideration of the implications of the agricultural revolution.

Courtesy Rainbird

Richard Leakey at work.

What, then, of the whole? Leakey himself comes over as a sympathetic and relaxed linkman, whose own personality does not obtrude between the viewer and the story, and the TV series is beautifully filmed by Alec Curtis and John Record. The scope of both the series and the book is satisfyingly broad and, as we shall see, there is a message of hope. The book is well written and illustrated, though I could not help feeling that a figure showing the relationships of the hominids to one another in time and their possible evolutionary relationships would have been helpful. Though there is a good index, there is no bibliography of any kind, not even to other popular books.

The culmination of Leakey's message comes in his consideration of the origins of human aggression. He suggests that it was only after the agricultural revolution, about 10,000 years ago, that the possession of settled and accumulated property made its gaining by aggression profitable and its defence therefore necessary. So, Leakey concludes, "Warfare is a social and political response to changed economic circumstances. Humans are essentially cultural creatures capable of responding in many different ways to the same prevailing circumstances". For him, the ills of the modern world are not the ultimate result of the irrevocable crystallization of our social nature during some early man-the-hunter phase of evolution, but are a comparatively recent adaptation — and therefore reversible. Let us hope that history will prove him right. □

Barry Cox is Professor of Zoology at King's College, University of London.

Glimpsed behind the curtain

Vera Rich

The Social Context of Soviet Science. Edited by L.L. Lubrano and S.G. Solomon. Pp.245. ISBN 0-8919-0712-9. (Dawson, Folkestone, Kent/Westview, Boulder, Colorado: 1980.) £10, $26.25.

ALTHOUGH the average Western scientist may well have numerous contacts with Soviet colleagues, and may himself have visited the Soviet Union, he will not necessarily have an overall view of the structure of Soviet scientific life. This is not for lack of information — as Susan Gross Solomon observes in her introductory essay, the Western study of Soviet science is now entering its fourth decade, with a proliferation of research projects addressed to governmental as well as academic audiences. That this information fails to reach those scientists who would benefit from a better understanding of the milieu in which their Soviet counterparts live and work, is perhaps due to its being published primarily in the sociological literature. The merit of the present work is not so much that it provides any fundamentally new material, but that it presents the consensus of present thinking in a readily available form.

This is not to say that the essays have been written with scientists in mind. As Loren R. Graham observes in "Reasons for Studying Soviet Science", the principal reason for compiling this book was the desire to understand better the nature of the Soviet Union as a whole. The emphasis given to the Lysenko affair (which, in Mark Adams's "Science, Ideology and Structure", occupies a major place in the book), or the controversy over the pollution of Lake Baikal of the mid 1960s, is thus not so much because of the scientific importance of the issues, but as examples, on the one hand of how political influence can effectively destroy a whole scientific discipline, and, on the other, of a rare case of public disagreement within Soviet society over a major policy issue. Indeed, at times this emphasis may prove enlightening to the scientist. As Thane Gustafson points out in his essay "Why doesn't Soviet Science do better than it does?", the simple answer which jumps to mind — "political interference" — is not necessarily the whole truth. "Crude ideological or political interference", he says, is the "exception not the rule"; the hierarchical structure and centralized planning system in the Soviet Union, the educational system which is stronger on theoretical than practical training and the uniting of administrative and research decision-making powers in the same individual must all be taken into account. Whether or not one wholly agrees with his conclusion that the "arrangements developed in the United States are more effective in creating a favorable environment for rapid scientific advance than those of the Soviet Union", his ideas are a useful counter-balance to émigré writers such as Mark Popovsky who attributes the Soviet short-fall in, say, Nobel prizes to political considerations.

Volume 3 of *True Visual Magnitude Star Atlas, Northern Stars,* by C. Papadopoulos and C. Scovil (for review see *Nature* **284**, 680; 1980) has been published by Pergamon.

How valid it is to judge scientific productivity by such methods is a moot point. Linda L. Lubrano in her essay on scientific collectives does not commit herself; she simply reproduces a number of Soviet tables, assessing productivity in terms of papers written and candidates trained, to illustrate the Soviet concern with norms and output. When she goes on to review Soviet studies on scientific motivation and the psychology of research workers, her exposition lapses into a turgidity and dullness typical of much of the Soviet writing on the subject.

Unfortunately, Professor Lubrano's article was written before the launching of the current press debate in *Literaturnaya Gazeta* on the prestige of pure, as opposed to applied research, many contributions to which have dealt with precisely the same problems of motivation and performance assessment in a far more lively form. This gives a slightly dated air to Lubrano's article, as — although to a lesser extent — it does to the two essays on applied research: Bruce Parrott's "The Organizational Environment of Applied Research" and Kendall E. Bailes's "The Technical Specialists' Social Composition and Attitudes". The latter is almost entirely sociological in content; the former, much wider in scope, ranges over such issues as the changing role of the Academy of Sciences in promoting applied research, the amalgamation of R&D units with production enterprises in the 1970s, and even the degree to which Soviet scientists are alienated from the bureaucratic system.

Taken individually, the seven essays cover a wide range of topics essential to a better understanding of Soviet science. Unfortunately, they remain seven essays, rather than becoming a book in their own right. There is no index, nor are there any useful aids such as a glossary of Soviet acronyms and organizations.

For a scientific audience, the main deficiency of the book is that it was not primarily intended for scientists. Yet, in the absence of a more suitable work, it is to be recommended as background reading to anyone who has dealings not only with individual Soviet scientists, but also with the Soviet scientific establishment. □

Vera Rich is a journalist who specializes in East European affairs.

Tumour as parasite

P.B. Medawar

The Interaction of Cancer and Host: Its Therapeutic Significance. By Michael F.A. Woodruff. Pp.467. ISBN 0-8089-1265-8. (Grune & Stratton/Academic: 1980.) $46.50, £26.20.

Sir Michael Woodruff, FRS, is one of a small and important group of surgeons that was brought into being by transplantation therapy: a man equally at home in ward and laboratory who has contributed to both the science and the arts of transplantation. The range of topics under discussion here and the depth of their treatment are such as we normally associate with symposium proceedings or multi-author volumes. Neither is a satisfactory genre for scientific communication, for multi-author works usually have a rootless and untidy character because each author busies himself with the part of the business entrusted to himself, often to the neglect of more general considerations.

The volume under review is open to no such reproaches. It is the work of a single mind — and a very good mind it is too: spaciously well informed, critical where criticism is called for (witness Woodruff's treatment of the so-called "autonomy" of tumours and the discussion of the notion of the immunological surveillance), but sanguine too in its appraisal of the contribution laboratory science may make to the prophylaxis or treatment of tumours.

A majestic list of references citing titles and first and last page numbers occupies pp.289–419. Over a long period of time such reference lists repay their authors for some of the time and trouble that goes into them but we readers are beneficiaries right away.

In the preparation of a work of this scale and unified structure the time must be set when copy is all packed up and sent to press: no useful purpose is served by delaying publication in the hope that someone will come forward with a revelation without which the book would be incomplete. Woodruff's work is conceptually up to date and its character is such that it matters not the least that it does not include every contribution to tumour biology up to the date of publication of this issue of *Nature*. No reference is made, for example, to the anti-tumour action of retinoids nor to recent epidemiological evidence pointing to the complicity of vitamin A in resistance to tumours. These very recent advances were known only as premonitory rumblings to those who were listening for them. □

Sir Peter Medawar is Head of the Transplantation Biology Section of the Medical Research Council's Clinical Research Centre and is currently engaged in research on tumour prophylaxis.

IQ: the environmentalist position

Richard Lynn

Race, IQ, and Jensen. By James R. Flynn. Pp.313. ISBN 0-7100-0651-9. (Routledge & Kegan Paul: 1980.) £10.50, $27.50.

Hitherto, the environmentalist case on the race and IQ issue has been characterized more by emotion and ideology than by reasoning and science. This deficiency is to some degree remedied by the present volume in which Dr Flynn gives a cogent and reasoned exposition of the environmentalist position.

The author recognizes that the chief problem for the environmentalist is to find an environmental factor capable of depressing the mean Negro IQ by one standard deviation or 15 IQ points. He recognizes also that many of the possible candidates frequently advanced by environmentalists do not begin to stand up to examination. This is particularly true of both poverty and discrimination. Among other problems these possibilities collapse in the light of the experience in the United States of the Jews and the Chinese. Both of these ethnic groups arrived as penniless immigrants and suffered initially from the handicaps of poverty and discrimination. Yet both minorities have mean IQs at least as high as those of Caucasians of north-west European descent.

It remains possible, however, that there is a more subtle environmental factor depressing the Negro IQ, or perhaps a set of factors each exerting a relatively small effect. The author proposes low parental IQ, poor nutrition during the mother's pregnancy, lack of stimulation in early childhood, family dislocation, low self-image and poor education. Taken together the author argues that these could be sufficient to account for the phenomenon. The difficulties with these arguments are twofold. First, as the posited deleterious effects are small they are hard to detect. Second, there is a chicken and egg problem in so far as the proposed factors are mainly functions of low IQ in Negro parents.

While there are manifest weaknesses in the environmentalist position, there are also problems in the genetic hypothesis. Two of these are particularly important. In the first place, the genetic hypothesis requires that in individuals of mixed Negro and Caucasian ancestry the IQ should be positively related to the proportion of Caucasian genes. There is contradictory evidence on this point and it is certainly by no means clear that this is the case. Secondly, there is a study of the children of German mothers and American Negro servicemen, in which it was found that the IQ of these children differed little from the mean of the German population.

It is a pity that the author has largely confined his discussion to the American data. The results of intelligence testing of racial and ethnic populations in other parts of the world are an important dimension of the problem and should be considered in any comprehensive discussion. Thus, for instance, if there were a Negro population somewhere in the world with a mean IQ of 100 the environmentalist case would be considerably strengthened. There is a substantial amount of evidence on this question from Negro populations in London, Jamaica and a number of countries in Africa, but in all cases their mean IQ is either about the same as that of American Negroes or else lower. These results cannot be readily explained in terms of test bias, because Mongoloid populations in Japan, Singapore, Taiwan and Hong Kong obtain about the same mean IQ as those of Caucasians in the United States, north-west Europe, Australia and New Zealand. This consistency of intelligence levels across many different nations would appear to suggest a significant genetic determination.

While the present volume is limited in so far as it does not come to terms with this important set of data, it has to be welcomed as a scholarly and informed presentation of the environmentalist position. □

Richard Lynn is Professor of Psychology at the New University of Ulster, Coleraine.

IRCS
JOURNALS

Keeping pace
with science

IRCS Medical Science
St. Leonard's House
St. Leonardgate,
Lancaster, U.K.
A member of Elsevier Science Publishers

Circle No.12 on Reader Enquiry Card.

Einstein refined — just in time

W.B. Bonnor

Exact Solutions of Einstein's Field Equations. By D. Kramer *et al.* Pp. 425. ISBN 0-521-23041-1. (Cambridge University Press: 1981.) £30, $72.

PROGRESS in theoretical physics consists largely of formulating equations and solving them. For instance, the equations put forward by Newton, and by Maxwell, are still being solved for application to special problems; and theorists have been seeking solutions of Einstein's equations since general relativity was discovered in 1915.

Solutions, yes — but why *exact* solutions? Aren't approximate solutions good enough? E. Schrödinger wrote (about Einstein's unified field theory) that "the exact solutions, involving strong fields, have disclosed the ingenuity of the mathematicians who discovered them, but nothing more". Approximate solutions are certainly good enough when a theory is well understood, like Newton's mechanics, and there is no danger of series failing to converge. But in a mathematically complicated theory like general relativity one does not always trust approximations. For example, after half a century of approximate work on motion, there is still no agreement on whether freely falling bodies radiate gravitational waves. A suitable exact solution could settle this once and for all.

A vast number of exact solutions of Einstein's equations have been discovered since 1915. With the half-dozen or so exceptions to be found in nearly all textbooks, these have lain buried in the literature. A warm welcome will therefore be given to this book. The magnitude of the authors' task is indicated by the bibliography which contains about 1,000 references; and these have been sifted from a list of twice as many. Most of the solutions given have been checked, which must have been a monumental task. Indeed, perhaps the work has been done only just in time: had it been left any longer it might not have been possible at all, because papers on exact solutions are appearing at the rate of over 100 per year.

It is possible to use the book simply as a catalogue in which to search for an exact solution, knowing to begin with some of its properties. However it is much more than this. It was essential to impress some order on the plethora of exact solutions obtained by a variety of methods. The authors choose as their principal methods of classification, first, symmetry groups and, second, Petrov types. They give an excellent introduction to both these topics, including much not easily found in other books. Part I, General Methods, also includes chapters on differential and Riemannian geometry, and on the Newman–Penrose formalism.

The authors have done such a fine job that it is important to realize what the book does *not* include. First, as stated in the Introduction, the exact solutions given are those arising from the following energy–momentum tensors: vacuum, electromagnetic fields, pure radiation, dust and perfect fluids. Although all the really important exact solutions fall under these headings, it should be noticed that some areas of physical interest are excluded, for example, charged matter and viscous fluids. It must also be remembered that nearly all solutions in the book are purely local, and no consideration is given to their global existence. Nor are boundary value problems within the authors' terms of reference, so the matching of different solutions is not considered; this leads to the somewhat bizarre omission of a famous solution, that of Einstein and Strauss for a star embedded in the expanding Universe. Another possible source of misunderstanding for the unwary reader lies in the quite forgiveable practice of giving the general form of a solution wherever possible, and omitting reference to special cases, even where these may be of physical importance.

All seriously interested in general relativity should get this book. They should also take advantage of the authors' invitation to inform them of omissions and of new solutions. In this way we may hope that the catalogue will be kept up to date and re-published from time to time. ☐

W.B. Bonnor is Professor of Mathematics at Queen Elizabeth College, University of London.

Les sciences bien vulgarisées

Alan Isaacs

Concise Encyclopedia of the Sciences. Edited by J.-D. Yule. Pp.590. ISBN 0-87196-491-0. (Facts on File: 1981.) $29.95.

BEING able to praise one's competitors is an indulgence not frequently available to those of us who earn a living by what the French call the *vulgarisation* of science. However, this is an excellent book, well written, expertly edited and lavishly illustrated mostly in full colour — or rather color, for this is the American edition of a book one version of which has already been published in the UK (Phaidon, 1978; £9.95). At $29.95 the American edition is undoubtedly good value — for the youngster with an interest in almost any branch of science it will certainly be a most acceptable gift. Uncles (and aunts) please note. It will also be a valuable asset for any library and for readers with no specialized knowledge of science.

Broadly speaking it does exactly what it says it will do — supplies both a dictionary of the most commonly encountered words of science as well as the encyclopaedic background material necessary for understanding their use in a wider context. The 400,000-word text is competent, accurate and jargon-free; for a book of this size it is difficult to see how it could be improved. One third of the text area is devoted to illustrations, almost all of which are a pleasure to look at, both because they communicate the information they are intended to communicate and because they are either well-chosen photographs or well-drawn diagrams. Particularly effective are the historical panels that summarize the history of some subjects and the "context panels" that illustrate the interrelationships between branches of other subjects.

If one has to criticize, one could say that it is perhaps unwise to have claimed (as the jacket does) to have covered philosophy and medicine. In both these fields the coverage is rather less than adequate. In philosophy, for example, the extent of the lacunae can be judged from the fact that there are no entries for ethics or existentialism, nor are Duns Scotus or Aquinas to be found. As Lady Bracknell might have said — to lose one scholastic doctor may be regarded as a misfortune; to lose two looks like carelessness!

In medicine the coverage is more thorough but it is perhaps surprising to find an entry for lymphoma but not one for the more common but less ominous papilloma. Some of the more recent techniques, such as tomography and thermography, are also omitted, as is the now ubiquitous Emiscanner.

If no attempt had been made to cover either philosophy or medicine there would perhaps have been room to deepen the treatment of the growing points of physical science. Quarks, for example, get only a one-sentence mention under subatomic particles, the state of play in nuclear fusion and fast-breeders is not explained, and the alternative energy field is not very well covered (nothing on wave or wind power or biomass energy).

It is, however, uncharitable to dwell on these gaps; so much is included and treated so well. This is undoubtedly an excellent reference book. ☐

Alan Isaacs is editor of the Macmillan Encyclopedia, *the* Penguin Dictionary of Science, Longman's New Dictionary of Physics, *and science editor of* Collins English Dictionary. *He has contributed to many other dictionaries and encyclopaedias.*

nature

14 May 1981

Will Mitterand be free to change?

What happens when one of the most conservative countries in Europe, perhaps even in the world, elects a socialist as a president? Revolution? Or nothing much? Or a mixture of the two? This is what people outside France are asking now that M. François Mitterand has beaten the incumbent president more handsomely in last Sunday's election than even the optimists of the left had been hoping. Frenchmen, however, know the truth.

The new president will be more tightly contained by economic and social constraints than his ardent supporters will be wishing. (M. Mitterand himself, on the other hand, may relish the lack of absolute freedom that he will discover when he takes over in ten days or so.) For the French economy, like others in Western Europe, is balanced on a narrow ledge separating the disasters of inflation and of slump. The temptation of any socialist government to cure unemployment by spending borrowed money is not a cure but a recipe for another kind of disaster. The new president will also find that even his most radical supporters are as firmly wedded to the preservation of familiar institutions as die-hard conservatives. This is why France in the 1980s is so obviously an amalgam of the recent past and the seventeenth century. The plastic bodied automobile on the once-Roman road now lined by poplars is as good a symbol as any.

The influence of the past quarter of a century, since the creation of the Fifth Republic, is nevertheless crucial both for France and for M. Mitterand. With the ending of the neo-Gaullist era, it will be tempting to forget how much has been accomplished. At the beginning of the period, French science and technology were not much to be proud of. There were a few nuclei of industrial innovation, the nuclear energy industry and the new plants for exploitation of natural gas and petroleum in the south and in North Africa, but the telephone system hardly functioned. University research was in the doldrums, chiefly from lack of funds even for teaching the huge numbers of students flocking to the universities. The CNRS was making modest experiments with its scheme for supporting members of staff within university departments, but neither the universities nor the council had begun to appreciate the importance of this device. The university system was hardly equipped to do so, as the upheavals of 1968 demonstrated all too clearly. Yet with the passage of time, French research has prospered. For some reason, imaginative laboratory work has flourished while the government machine has been occupied with the design of frequently artificial programmes built around specified technological goals — the building of an aircraft, or the exploitation of technologies with made-up names, *l'informatique,* for example. Now and again, this peculiarly French system of dual support for research runs into trouble (see page 101). But on the whole, it has functioned better than could have been expected.

How will M. Mitterand change all this? Nobody at this stage can be sure. Part of the explanation for the new president's success is that he has kept his cards close to his chest during the election campaign. And some of the things he has been saying no doubt owe much to the need to appeal to particular groups of electors. Thus the promise that there will be a "pause" during which the future of the previously ambitious French nuclear power is re-examined may have been designed to win round the ecologists, who captured four per cent of the vote in the first round of the elections. The political cost of reviewing the plans for building nuclear power stations (and for the development of the fast reactor called Super-Phénix) would not be high. Deciding to cut them back would not merely offend a large number of French workers but would run headlong into conflict with the traditional French conviction that self-sufficiency comes above prudence, economic or otherwise. If the doctrine of self-sufficiency is to be challenged, there are other fields in which the battle could be fought more profitably.

The most obvious opportunity for change is in the not entirely unrelated field of arms control. Since the collapse of the old scheme for a European Defence Force in the mid-1950s, France has been the stumbling block in a long series of arms control negotiations. Once President de Gaulle had accepted the arguments of his generals, General Gallois in particular, that national security required self-sufficiency in nuclear weapons, France has stood aside from sensible arrangements such as the Non-Proliferation Treaty and the Partial Test-Ban. In recent years, French policy has been obviously ambivalent. Thus France has resolutely refused to accede to the Non-Proliferation Treaty while behaving — in dealings with India, for example — as if it had done so. A moderate Gaullist government such as that of President Giscard d'Estaing has inevitably been the prisoner of the true Gaullists to its right, and thus politically prevented from throwing over the foolish doctrines of the 1950s. M. Mitterand will have no such inhibitions. On the contrary, his policies towards the world outside France predicate change. We shall soon find out how resolutely he seeks to amend the French position on nuclear self-sufficiency — and whether in the process he discovers that after a generation, the doctrine of comprehensive military self-sufficiency goes deeper than merely Gaullism.

The prospect that there will now also be a further acceleration of the growth of public spending on research and development is less secure, if only because there are both financial and structural limitations on what the new government can hope realistically to accomplish. French expenditure on research and development, by government and private industry, now amounts to 1.7 per cent of the Gross National Product. M. Mitterand has promised that this will increase by more than a half in the next five years. He has not explained how. And he would find it hard to do so even in the post-election euphoria. Channelling close on an extra one per cent of the national wealth through the government machine, earmarking it for spending on research and development, would be a formidable task if there were a decade in which to do the trick. To attempt it in five years would send inflationary ripples about the place. When he and his advisors start in earnest planning how to implement this promise, they will be driven to conclude that what is true elsewhere is also true in France — that the only basis for spending more money profitably is to begin by improving the quality, and increasing the scale of, university training in science, technology and research. And changing universities is slow work.

The proposal that there should be a minister of cabinet rank with responsibility for research and development is, on the other hand, more than merely sensible. In the circumstances that the new president will inherit, the most immediate danger is that organizations conceived of by bureaucrats and given life by politicians will become so powerful and set in their ways that they are at once wayward and uncontrollable. For all its virtues, the CNRS is in danger of becoming such a law unto itself. The practice of putting university researchers on the payroll, necessary though it may have been a generation ago, now threatens the autonomy of universities and impairs the flexibility of the research enterprise. The first task for the holder of this post (if and when appointed) will be to look again at the relationship between the universities and the rest of publicly supported

0028-0836/81/200097-02$01.00

research and development. In doing so, he or she will find it necessary to ask whether the condition of the French university system is entirely suited to present needs. In the past seven years of the presidency now ending, much has been done pragmatically to improve the lot of the universities, but very little is written down on paper. The new president seems to have been offering the nearest that France would accept to an American way of administering science. The question still to be answered is whether he will discover, in this connection too, that tradition goes deeper than voting behaviour.

Calculating flop

The British government, always ready to seek to demonstrate that water can be made to run uphill by a sufficient display of determination, has been in trouble for several weeks because it insists on willing that Britain, like the United States, should enjoy the benefits of an indigenous, productive, competitive and profitable mainframe computer manufacturer. The trouble came to a head on Monday, when the chairman and managing director of the company called International Computers Limited (ICL for short) resigned, after the Department of Industry had put a pistol to their heads and asked that they should accept a new way of working. Mr Phillip Chappell, the part-time chairman for just over a year, will go back to being a full-time merchant banker. His managing director, Dr Christopher Wilson, will remain a director of ICL (and will be paid an executive director's salary). Mr Chappell and Dr Wilson are replaced by two outsiders — Mr C. C. F. Laidlaw, vice-chairman of British Petroleum Ltd, as chairman and Mr Robert Wilmot (now managing director of the British end of Texas Instruments) as the new managing director. The government's headhunters have done well; the new management has at least as good a chance of succeeding as the old. But the chances are not high, and it is in the (British) public interest that everybody concerned should acknowledge that to be the case.

ICL has a long distinguished history, going back to the end of the Second World War. That was when people straight out of the design of radar systems settled down to making electronic versions of the old electromechanical calculating "engines". For a time, at least in Britain, it seemed that success exceeded all reasonable expectations. A private company called Ferranti was closely involved with the development of a computing system at the University of Manchester, linked with the names of people such as Thuring and F. C. Williams. The first prototype of this computer was frequently inspected by admiring visitors from other computer companies, many of them from overseas. In due course, when it seemed likely that an electronic computing machine of some kind might sell and make money, the then British government negotiated with Ferranti an arrangement under which there would be a public subsidy to help put this new machine on the market. The amount concerned, just over £1 million, was tiny in comparison with the task. But Ferranti chose to spend this pittance not on the development of a commercial market but on the development of a more advanced and less successful machine.

With the passage of the decades (almost exactly two of them), it became plain that the promise of the immediate post-war period had not been fulfilled. By 1968, the turnover of IBM World Trade (the overseas arm of IBM Inc.) probably exceeded the annual exports of the United Kingdom. Enthused by a vision of a brave new technical world, however, Mr (now Sir) Harold Wilson, then the British Prime Minister, set up an organization called the Industrial Reorganisation Corporation to effect the merger of companies judged too small to compete on their own with the giants overseas. In 1968, Ferranti's computer interests, together with those of the other British manufacturers eager to build mainframe computers, were welded together with the help of public money. The stage was set for the emergence of a giant to complete with IBM *et al.* The snag was that nobody thought, knew or seemed to care what kind of company ICL should be so long as it was big.

This week's announcement that the management must be changed yet again is therefore a sign that water is still steadfastly refusing to flow uphill. The attempts by successive British governments to give ICL an inside track in the supply of computers for the public sector has most often been defeated by the preference of the intended customers for somebody else's machines. ICL shares with the most successful of its overseas competitors, even IBM, the fault of not having judged just when small computers would come into their own. Unlike some of them, it has been too insecure to ride this tide. Now, it seems, ICL is hoping to make up with its flair for distributed computer networks for its failure to make much money on the huge machines that are its pride and joy. British taxpayers in particular will wish it luck. In the next two years, the company will survive with the help of a loan of £270 million from the commercial banks, £200 million of which is underwritten by the government on behalf of the taxpayers.

In retrospect, what went wrong at ICL was not the decision (in 1968) that there should be a unified computer company, nor the British government's repeated reappointment of new managers, but the failure to acknowledge at the outset that in a novel field of technology, it is impossible to design by means of a committee the commercial enterprise that is certain to carry off the prize. The best that can be hoped for is to send several arrows at the target.

The irony in all this is that British government's stake in ICL is now much less than it used to be a year ago. Technically, the company is much more directly under the control of its private shareholders than of the Department of Industry since the government and some other big shareholders sold out to the general public at the top of the market. But ICL will remain in hock to the department so long as the £40 million public loan arranged some years ago remains unpaid. For the whole of that time, British governments will find it necessary to worry about the strategic value of companies such as ICL. Some of them may be able to stomach pulling out of old industries such as shipbuilding, but pulling out of computer manufacturing may seem like turning the back on the future. In ICL's present crisis, all kinds of interests are seeking to persuade the British government of that truth. But it is not a truth at all, of course. Consistently unprofitable companies serve no purpose and should be put to sleep, whatever they happen to be making.

Tokenism again

Public officials, who necessarily spend their waking hours in committee meetings, naturally believe that appearance is reality. For the past several years, but especially since the appearance of the Finniston Report on engineering education eighteen months ago, the British government and the whole company of its officials have been dedicated to the belief that the status of engineering as a profession and of engineers as people must somehow be enhanced. Otherwise, the argument goes, British manufacturing industry will continue to be uncompetitive and will go out of business.

Dr Duncan Davies, until last month Chief Scientist and Engineer at the Department of Industry, and Sir Geoffrey Allen, the chairman of what used to be the Science Research Council, have shown what can be done. Dr Davies has rechristened himself "Chief Engineer and Scientist". Allen's research council, with the help of Queen Elizabeth II, has been renamed the Science and Engineering Research Council. Everybody concerned proclaims that the change of name does not mean a change of function. But engineers (as distinct from scientists) will walk a little taller. And the only cost will be that charged by the suppliers of printed stationery to public agencies and corporations. Nobody seems alarmed that the chief recommendation of the Finniston Report — that there should be an Engineering Council to regulate the profession of engineer — has still not been set up.

At a meeting on Thursday this week, the unions claiming to represent professional engineers will be throwing yet more sand in the works. Tokenism, not the best policy, may be the only one.

0028-0836/81/200098-01$01.00

Congress will battle against proliferation

Senate split confirmation of Malone

Washington

Democrats in the US Congress are shaping up for what is expected to be a major confrontation with the new Republican Administration over proposals to loosen controls on the export of nuclear technology and materials. Already they have tried, albeit unsuccessfully, to block Mr Reagan's nominee for the position of Chief Nuclear Advisor to the State Department — Mr James Malone — as a warning shot to the Administration that it can expect fierce opposition if it tries to relax existing controls too far or too fast.

Mr Malone had previously been the General Counsel of the Arms Control and Disarmament Agency (ACDA) under President Nixon. Shortly after Mr Reagan's election last November he was appointed to head the transition team for ACDA. And his nomination for the State Department post, officially known as Assistant Secretary for Oceans and International Environmental and Scientific Affairs, has been warmly supported by the conservative Republican, Senator Jesse Helms of North Carolina.

Last week his nomination was formally approved by the Republican-dominated Foreign Relations Committee in the Senate, and confirmation by the full Senate remains a formality. However, five of the fifteen members of the committee voting cast their ballots against Mr Malone, expressing concern that the type of proposals he was supporting for speeding up nuclear exports would increase the risk of the proliferation of nuclear weapons.

Several Democrat members of the Senate committee, including Senator John Glenn of Ohio, a strong supporter of stiff anti-proliferation measures, also raised concerns about possible conflicts of interest. During the Carter Administration Mr Malone has been a member of a Washington law firm, Doub and Muntzing, which includes among its clients groups in Japan who have been seeking permission to export nuclear technology from the United States.

During the committee hearings on his nomination, Mr Malone was reluctant to provide much detail of the extent to which his company had been involved in furthering the goals of such clients. His reluctance provided Mr Glenn with an excuse to hold up his nomination for several weeks, and the committee members unhappy with his appointment are proposing strict procedures to prevent any conflict of interest.

Their main concern, however, is a broader one, namely, that Mr Malone seems keen to dismantle many of the nuclear export controls introduced during the Carter Administration, in particular parts of the Nuclear Non-Proliferation Act of 1978. This act requires countries to accept international safeguards on their nuclear installations before they can receive US-produced nuclear technology or materials, and to obtain US permission before these are transferred to a third party, for example, for the reprocessing of spent fuel.

The transition team report on ACDA suggested that a new office be created within the State Department that would centralize all aspects of nuclear export policy — including the assessment of whether the receiving country was obeying international safeguards, a role at present performed by the Nuclear Regulatory Committee. The nuclear industry, already facing the prospect of a major slump in domestic demand, favours such centralization in order to expedite the licensing of exports; but Senator Glenn and others reply that the Nuclear Regulatory Committee (NRC) was given control of export licensing specifically so that decisions on safeguards would be made on technical rather than political grounds (as happened when NRC refused to allow to the export of nuclear fuel to the Tarapur nuclear plant in India, a decision later overruled by President Carter).

The nuclear industry is hoping for more sympathetic support for its international trade from the new Administration than from its predecessor. The Atomic Industry Forum points out that Secretary of State Alexander Haig has already called nuclear energy "imperative" for Western industrialized nations, and also notes that the appointment of ex-NRC commissioner Richard Kennedy as Under-Secretary of State for Management places "a nuclear supporter in Haig's inner circle" — even though the post has no nuclear responsibility. The industry would welcome a return to the "reliable supplier" doctrine as a device for limiting the misuse of nuclear technology which Mr Malone is expected to promote, but which critics claim has in the past proved ineffective in preventing nuclear exports claimed necessary for energy programmes being used for military purposes.

Ironically, the major threat may come from a very different direction, the Administration's desire to reduce public spending substantially. Among the agencies expecting major budget cuts is the Export-Import Bank, which provides loans to US companies to subsidize their exports, particularly to Third World countries. Virtually all US nuclear exports have been paid for with loans from this bank. If the source of the loans dries up — as seems possible — the nuclear industry will find itself facing a new set of hurdles to cross.

David Dickson

Science Adviser appointment revived

Washington

The appointment of a Presidential Science Advisor, on the back burner for most of the time since Inauguration Day, has become a live issue again. In the past few days, the name of a potential candidate for the job has been circulating. Earlier this week, it seemed certain that the new man would be formally nominated before the end of May.

The newest candidate is Dr George A. Keyworth, 41-year-old head of the physics division at the Los Alamos National Laboratory in New Mexico, where he has worked for the past 13 years in nuclear weapons research, laser fusion and basic nuclear physics.

Dr Keyworth has already visited Washington to discuss the job with White House officials, but said last week that he has not formally been offered the position — although the offer was widely expected in Washington. He holds a bachelor's degree from Yale University in 1963, and a PhD in physics from Duke University in 1968.

Dr Keyworth is largely unknown in the scientific community; he seems to have had little experience in Washington, although he has served on the Department of Energy's Fusion Data Committee. Dr Keyworth's name is said to have been put forward by the domestic policy staff within the White House, and he is said to have the enthusiastic support of Dr Edward Teller, one of Mr Reagan's advisers when he was governor of California.

The Reagan Administration has apparently had considerable difficulty in finding a candidate prepared to accept the post of science adviser and director of the Office of Science and Technology Policy, partly because of rumours that the role of the office is likely to be downgraded from its prominence during the Carter Administration under Dr Frank Press. Dr Keyworth says that he is in complete agreement with the new Administration's attitude towards science, and attempts to restore "American pre-eminence in industrial technology".

Dr Nunzio J. Palladino, dean of the College of Engineering at Pennsylvania State University, is expected to be named by Mr Reagan as the new chairman of the Nuclear Regulatory Commission. Dr Palladino worked as an engineer with the Westinghouse Electric Company in Philadelphia before joining the National Laboratory at Oak Ridge.

David Dickson

0028-0836/81/200099-01$01.00

French presidential election

Good for science?

Champagne corks were popping in science departments throughout France on Monday at the news that François Mitterand, socialist party leader, was to be France's next President. Not that most French scientists are socialists; rather, that science will take a leap forward in France if Mitterand sticks to his announced policies.

Mitterand has promised a "national council for science and culture", reporting to the president, and has announced that spending on research and development in France should reach 2.5 per cent of gross national product by 1985, compared with the present 1.8 per cent. Mitterand's goal could only be reached with high government spending, for French industry is notably niggardly in its provision for science; Giscard had a more cautious target of 2.3 per cent by 1988.

The socialist leader has also said he will raise the rank of the minister of science to full cabinet level, so the minister will participate in all major policy decisions. (The current incumbent, Pierre Aigrain, attends cabinet only when matters involving science are under discussion.) Two names are being tipped for the job; François Gros, a biochemist at the Institut Pasteur, who has been active among scientists on Mitterand's behalf; and the career politician Jean-Pierre Chevenement, a member of the chamber of deputies who has taken an interest in scientific affairs. The high level of the post may favour the politician.

Mitterand will also set up a parliamentary special committee for technology assessment, something previously unheard of in France, where technological decisions have been reached in an entirely autocratic manner. The new president is also cool on nuclear power, so France's nuclear lead might be expected to slip unless national pride erodes the president's doubts. He has said he will call for a "pause" in nuclear construction after the reactors now under order, and the fast breeder Super-Phénix, are completed, in order to review the programme. **Robert Walgate**

French oceanography

Pulling together

The outgoing French government completed the last leg of its reorganization of oceanographic research last month with the creation of a joint interdisciplinary programme between CNRS, which is the main funding body for scientific research, and the Ministry of Universities. The new programme is expected to be in full operation early next year. The reorganization, which began at the end of 1979 when an informal Comité d'Océanologie was set up to advise government, reflects the political importance in France of technologies for exploiting the mineral and biological wealth of the oceans.

This attempt to coordinate research supportd by CNRS and the Ministry of Universities through the new Programme Interdisciplinaire de Recherche Océanographique (PIRO) is unusual in French research administration, for with the growth of support for oceanography since the Second World War, two disparate research programmes have grown up within the universities and the CNRS laboratories. Although similar difficulties have arisen in other fields, oceanographic research has been particularly affected, according to the director of PIRO, Roger Chesselet.

PIRO's task of coordination will fall to an eighteen-member scientific committee with an annual budget of 11 million FFr, the Ministry of Universities and CNRS each contributing half. The budget, which will be in addition to the 185 million FFr already spent on oceanographic research, will be used to run and buy equipment for six small research vessels, and to foster collaboration between several research groups working on different aspects of the same problem. PIRO, whose members are CNRS administrators and oceanographers nominated by CNRS and the Ministry of Universities, will be keen to support fundamental research in physical and biological oceanography, leaving mineral exploration and the application of new technologies to the well-established Centre National pour l'Exploitation des Océans (CNEXO). A small amount of the extra 11 million FFr will, however, be spent on research equipment for use by CNRS and university staff on board CNEXO's large, ocean-going research vessels.

The setting up of PIRO follows the definition of three categories of research: research with short-term application supported by the departments of industry, transport and environment; research with medium- and long-term application supported by CNEXO; and fundamental research, the responsibility of the Ministry of Universities.

Under the new arrangements, the different organizations supporting oceanography will report to the Comité d'Océanologie. France is also hoping to collaborate internationally, especially with developing countries, on pollution problems and the exploitation of ocean resources when the Law of the Sea becomes operational. **Judy Redfearn**

Bazan protests

The academic research community in Argentina has been disconcerted by the precipitate sacking of Professor Nicolas Bazan from his posts as Director of the Institute of Biochemical Research at Bahia Blanca and as Professor of Biological Chemistry at the University of the South. Professor Bazan (38) heard of his dismissal only when he returned to Argentina from the United States early in March, but it is held to have taken effect from the end of 1980.

The most puzzling aspect of this development is that Professor Bazan has insisted that he has no connections with political organizations deemed subversive by the military regime. One explanation is that his success has made him the target for influential jealousy, another that his dismissal is intended as a reminder that the influence of the military persists in the universities of Argentina.

Dr Bazan is a neurochemist who has, in the past few months, been able to demonstrate that in rat models of epilepsy, arachidonic acid and other precursors of prostaglandins are released into the brain in large amounts. Professor Bazan had been planning an investigation of the presence of arachidonic acid in the cerebrospinal fluid of human patients suffering from epilepsy, on the assumption that the underlying cause is a membrane deficiency.

Since his dismissal, Professor Bazan has insisted that he will fight for reinstatement. He has been supported by the staff of his institute, which sent a letter of protest to the President of Argentina on 21 March. The Argentinian press has also covered Bazan's dismissal sympathetically.

UK university finance

Truth looms

British universities should know more clearly next week what the future holds. For then the University Grants Committee will put out a circular spelling out the basis on which recurrent grants will be distributed to individual universities in the next few years. This first circular will be followed in June by more detailed information about the proposed division of the recurrent grant for 1981–82 among individual universities.

The committee has been forced to this two-stage approach by the need for detailed consultation with the Department of Education and Science about ways of sweetening what must certainly be a bitter pill. Originally it had been intended that final allocations would be announced this month. One of the concessions wrung from the government is likely to be an assurance that recurrent funds for the universities will be held for the next three years at the levels suggested in the white paper on public expenditure published in March.

British universities may thus know where they stand more clearly than in recent years, when government support has been determined one year at a time. The promised grants will, in the government's opinion, nevertheless require many universities "to review the range and

0028-0836/81/200100-02$01.00

nature of their contribution."

Next week's letter from the grants committee is expected not to go so far as to suggest that some universities should close. The emphasis, instead, will be on financial economies within existing institutions, with the recommendation that universities should think of closing departments which, for one reason or another, do not meet tests of efficiency and effectiveness. Universities and those who work in them will be especially interested in whatever arrangements the committee suggests for dispensing with unwanted members of teaching staffs.

Dr Edward Parkes, chairman of the University Grants Committee, has, however, made it clear in his evidence to a parliamentary committee that the impending cuts will not be made uniformly across the board. Universities will also be affected differently by the still uncertain enrolment of overseas students for the coming academic year. In its *Statistical Bulletin* for April, the Department of Education and Science shows that overseas student numbers declined by 7 per cent between 1979–80 and 1980–81.

Enrolments between the same two years declined by 10 per cent. Although the universities as a whole have more successfully maintained overseas student numbers than other institutions of higher education, some among them are in serious trouble, while all fear that new enrolments next October will decline further.

One part of the government's package of support for the universities has, however, been decided. The equipment grant for British universities will be increased by about 11 per cent to £76.7 million for the academic year 1981–82.

Hungarian research

Cost of elitism

Budapest

Budapest's Eotvos Lorand University is to have a new science building — at a time when the Hungarian government is drastically reducing spending in virtually every sector of the economy.

This is not the first time that the university has been promised such a building, which would unite its 44 small science "chairs" under one roof into four large departments — mathematics and physics, chemistry, life sciences and Earth sciences. Yet, in spite of the economic climate, the academics concerned are confident that this time the plans will come to fruition, because the reorganization can be justified as cost-effective. With all laboratories in close proximity, it will be possible for the various departments to share expensive equipment.

More significant in precipitating the decision, however, was the public debate on the role of the university, begun in May 1979 by the journal *Magyar Tudomany,* in which the reorganization of universities into large departments was favoured.

One of the government's aims is to break the tradition of academic elitism. Young scientists are reluctant to leave cosy academic niches in research institutes to enter the harsher world of industry and the universities. Although the reorganization will be welcome, any real moves to strengthen the role of the universities in research would first have to deal with one of the major grievances of Hungarian scientists, the difficulty of purchasing equipment and reagents from abroad.

Because of the centralization of foreign trade, such purchases have to go through a single import–export company, which means a delay of at least six months for every order. Hard currency is not the only problem — the same delay occurs with purchases from the Comecon bloc. Only in research projects with a government priority rating — such as work on genetic engineering done under contract with the chemical industry — can the purchasing process be accelerated.

Whenever this subject is raised with officialdom, scientists say, the response is always sympathetic — but without visible result. Until the matter is resolved, however, young scientists will have one more reason for seeking a long-term post in one of the Academy of Science's numerous research institutes. For the academy has its own purchasing agency, "Akadimport", and although the delays involved are still of the order of two months, this still gives such researchers a considerable edge.

Vera Rich

Satellite-linked computers

Network planned

Several different groups in Britain are cooperating in an experiment to link computers via satellite. The idea is to demonstrate to home and overseas customers how British technology can help industry and research institutions with their communications problems. Project Universe, as the experiment is called, is being supported to the tune of £3 million for three years by the Department of Industry, the Science and Engineering Research Council, British Telecom and two private companies, GEC-Marconi and the software house Logica.

The experiment is mainly to demonstrate how groups of users such as universities and companies with offices in different centres can communicate with each other through networks of computers linked by land and satellite. Six laboratories will house rings of computers, based on a network designed at the University of Cambridge, and will be equipped with an earth-station consisting of a 3-m diameter dish, a 14-GHz radio transmitter and an 11-GHz receiver. Communications within a laboratory will be via the land system and between laboratories via one channel of the European Space Agency's Orbital Test Satellite. Each computer will be able to communicate with any other in the system initially at a rate of 1 Megabit per second.

The Department of Industry, which is providing the largest sum of money, hopes that the project will help develop technology and demonstrate uses for the network.

The two major demonstrations of the network will be in transmitting large volumes of scientific and industrial data between centres. The Rutherford Laboratory is particularly keen to use the network to transmit mask design for large-scale integrated circuits, developed with its electron beam lithography machine, to universities and industry.

The project is expected to be in operation early next year. It will run for two years throughout the rest of the life of the Orbital Test Satellite.

Judy Redfearn

Research management

Audits go on

Despite the tangle of legal threats, and claim and counter-claim over their previous — and still unpublished — work on Sir Bernard Lovell's radioastronomy at Jodrell Bank, a group of the Science Policy Research Unit (SPRU) at the University of Sussex has been given the go-ahead for its next study: CERN at Geneva.

The group is headed by Keith Pavitt, acting deputy director of SPRU, with principal researchers John Irvine and Ben Martin, who have developed what they call the method of "converging partial indicators" to assess the relative merits of scientific research groups.

Inevitably, this is dangerous territory, normally confined to the secrecy of peer review panels and referees' reports, and the Sussex group's serious attempt to arrive at objective criteria raised hackles as soon as it was applied to real cases, which included a comparison of radioastronomy groups in Britain and abroad and a study of British optical astronomy.

All these studies are now essentially complete, but none has been officially published — largely due to the Sussex group's procedure of checking its figures with the relevant laboratories first. Predictably, laboratories whose work is implicitly criticized object to either the figures or the conclusions.

The SPRU group, however, has been unable to consider such questions because at an early stage it was denied access to research council files under the "40 year rule", which controls the availability of government documents — even though the research councils are not government departments. It therefore could not assess the conditions under which one group, say, received a major grant, while another did not.

Nevertheless, what counts in the end is results, and the SPRU group claims to have a method which gives an objective assess-

0028-0836/81/200101-02$01.00

ment of those results by using the relevant research community's own judgements of scientific value, both directly (through confidential questionnaires) and indirectly (through publication rates and citation frequencies).

A preliminary version of the radioastronomy study was read last year at a closed session of the Organization for Economic Cooperation and Development, and was very well received; but the wide circulation of that paper, particularly in the Netherlands and West Germany, whose own radio telescopes were also assessed (not all favourably) led to recrimination from Jodrell Bank, which did not do so well in the study. And now North-Holland, the publisher of *Research Policy*, in which the final paper was to appear, is said to be refusing to publish the radioastronomy work.

Moreover, the originator of the SPRU study, Sussex physicist Dr Norman Dombey, wishes to have no part in such publication, and says he disagrees completely with the work, mostly done by Martin and Irvine, and its conclusions. Dombey, although a high energy physicist, will not be a member of the group for its study of CERN, and argued with the Social Science Research Council that the CERN work should not be funded. However, after nearly a year's deliberations, the council has given the group a grant for 15 months' work on CERN and its future project LEP, but has insisted this time on extremely strict rules on confidentiality.

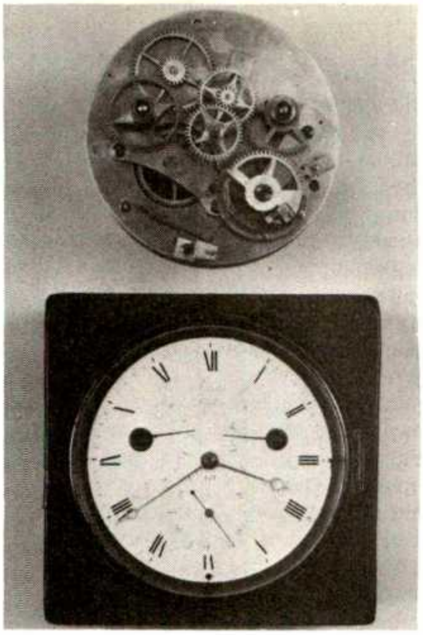
Robert Walgate

Arnold at Christie's

An 8-day marine timekeeper which is being offered for sale on 3 June at Christie's London auction rooms is thought to be a prototype for the modern marine chronometer, and it has been suggested that the spring detent which it employs could be the earliest ever, predating the putative "Earnshaw" pattern by several years. Made by John Arnold in London around 1775, the design of this un-numbered chronometer throws light on the many problems which confronted Arnold in his efforts to construct a timekeeper which would be accurate and reliable and could also be made in quantity at a relatively low price.

His chronometer (a term he himself coined) was designed to keep the same rate of going in every position and it incorporates a mechanism to compensate for changes in temperature. It has a three-armed uncut brass balance 45.2 mm in diameter. Regulation is by flat spiral balance spring, regulated by curb pins carried on a bi-metallic compensation curb, concentric with the balance pivot. The convex enamelled dial bears the signature "Arnold, London". His earliest dials were enamelled, but once commercial production was started, he changed over to silvered dials. The box, 6½ inches square and 3½ inches deep, is of mahogany.

Arnold first started working on marine timekeepers in 1786, about nine years after John Harrison — the carpenter turned clockmaker — had submitted his No.4 timekeeper to the Board of Longitude. This was a body which had been set up by the British government and which in 1714 offered a prize of £20,000 for a machine which would enable longitude at sea to be determined to a specified degree of accuracy. Harrison richly deserved the prize but the board shilly-shallied for many years, and it was only towards the end of his life that he recieved the final instalment of the award.



Had the Board of Longitude a free choice between the Harrison and Arnold timekeepers, the Arnold might well have been chosen on the grounds that it was less complicated and consequently easier to produce. Arnold made chronometers for the Admiralty at 60 guineas a piece, much cheaper than Harrison's which took about 3 years to make.

Arnold deserves much of the credit for introducing into England the manufacture of chronometers on a large scale.

As for the price likely to be reached when the chronometer is auctioned, Christie's are not willing to speculate. But if the device's history can be authenticated it could fetch anything between £20,000 and £100,000.

Arthur Frank

Mr Arthur Frank, the now-retired owner of a Glasgow-based optical business, is a well known collector of optical and scientific instruments. Much of his collection, which includes early telescopes, cameras and eye-glasses, is now distributed among several British museums. Mr Frank will be contributing occasional articles when important scientific instruments change hands or come onto the auction markets.

Indian government policy

Committed to science

New Delhi

The poor coordination between Indian scientists and politicians should improve following long-awaited decisions announced by Mrs Gandhi recently. Mrs Gandhi herself is to preside over a nine-member Cabinet Committee on Science and Technology, dealing with broad policy matters, and two more science committees have also been formed.

India's worsening position on the energy front and a massive bill for oil imports have led to the formation of a six-member Alternative Energy Commission, along the lines of the Atomic Energy Commission, to be chaired by Professor M. G. K. Menon, Secretary of the Department of Science and Technology. The third committee, the Science Advisory Committee to the Cabinet, consists of 20 eminent Indian scientists from all the major disciplines. Dr M. S. Swaminathan heads this body which is to advise the cabinet on science and technology policy.

The Cabinet Committee on Science and Technology will have an important role in deciding how to spend the science and technology budget of more than 30,000 million rupees (£1,667 million) allotted in the country's sixth five-year development plan. Main areas of interest will be in life sciences, oceanography, alternative energy sources and bioengineering as applied to pharmaceuticals.

Naresh Sahajpal

● **Energy alternatives:** The Indian government has given fresh impetus to the development of new and renewable sources of energy. The key point of the plan is reduced dependence on imported oil. The government is investing in systems for utilization of energy from the Sun, wind and biomass.

In photovoltaics, the public sector company, Central Electronics Limited (CEL) of Sahibabad, has begun bench-scale manufacture of single crystal silicon solar cells and panels with capacities of up to 10kW per year. Solar modules developed by CEL are already being used in powering a light beacon at Dwarka port, pumping water for a whole village in Rajasthan and lighting a 20-bed hospital in Ladhak. Solar water heating systems are being installed in many hotels, and a textile mill in Ahmedabad is already meeting its hot water requirements through a large solar heater.

A 120 million rupee (£7 million) project has been drawn up for developing low cost solar-grade silicon material and cheap fabrication techniques and for improving the efficacy of solar cells and solar panels.

A national wind energy centre is being set up at Allahabad Polytechnic to develop a low-cost windmill. A cold storage plant utilizing geothermal energy is being set up in Himachal Pradesh. And a 600 MW tidal power station is planned in the Gulf of Kutch in Gujarat.

Zaka Imam

0028-0836/81/200102-01$01.00

Third World development

UN funds falter

Washington

Frustrated in the United Nations by lack of progress in following up proposals from the Conference on Science and Technology for Development (UNCSTD), held in Vienna in August 1979, a group of developing countries has organized a top-level mission to try to persuade the oil-rich nations that they should be doing more to support technological efforts in the Third World.

In an initiative originating from government officials from Tunisia, Mexico, Pakistan and Guinea, a delegation made up of ministers or other top officials from ten developing countries will visit several major oil-producing countries at the beginning of next month to discuss ways of supporting technical assistance programmes. The countries to be visited include Saudi Arabia, Kuwait, the United Arab Emirates, Nigeria and Venezuela.

The delegation will then move to Paris, where it will report to the Secretary-General of the Organization for Economic Co-operation and Development (OECD). Several OECD countries, in particular the Scandinavians, have already expressed enthusiastic support for the so-called "Tunis initiative", which they see as avoiding some of the road-blocks faced by classical negotiating patterns on North/South economic issues.

Eventually it is hoped that a package can be put together in a form likely to be endorsed by the planned summit meeting between industrial and developing countries in Mexico City in October. President Ronald Reagan has recently agreed to attend the meeting (partly, it is rumoured, because President Lopez Portillo of Mexico has agreed not to invite Cuba's Fidel Castro).

The package could contain a commitment from oil-producing and developed countries to support technological development in the Third World by several hundred million dollars a year by the mid-1980s. The aim is to find a way of using the large cash surpluses being built up by the oil-producing nations to catalyse development efforts, without running into the political tensions which have polarized discussions between the Group of 77 and the industrialized nations within the United Nations.

In the event, the Interim Fund for Science and Technology, set up at the Vienna conference with an initial goal of raising $250 million to cover two years of operation, has fallen far short of its target. Less than $60 million has been received in pledges — and only about a third of this is thought to have been handed over. The interim fund is the responsibility of the United Nations Development Programme, and officials are working closely with the Tunisian government and other members of the so-called group of 10 — in addition to the four countries listed above, the other members are Peru, Senegal, Sudan, Kenya, Jordan and Sri Lanka — to find other ways of financing the projects that they have been assessing.

Although financial aid will be high on the agenda of next month's discussion with the oil-producing nations, equal emphasis is being placed on the need to support science and technology as an integral component of economic development in Third World nations. Even the oil-producing nations themselves are putting increasing emphasis on their need for technical assistance in building an economy which is not totally dependent on oil exports.

Officials from the United Nations Development Programme are also trying to link the efforts which have already been made under the interim fund to any plans emerging from the UN Conference on New and Renewable Sources of Energy, due to take place in Nairobi in August, for encouraging research on new energy sources. They argue that it would avoid unnecessary duplication if any new research initiatives from the Nairobi conference were under the same umbrella.

A second list of projects, totalling $11.2 million, has recently been approved for funding under the interim fund. These include $1.3 million to Brazil aimed at improving carbon fibre technology; $500,000 to Ethiopia for the popularization of science and technology; $670,000 to establish a unit for research on cellulose products in the Sudan; $370,000 for "interregional" courses in Bangladesh and Ghana on the application of modern techniques in physics to development (one on monsoon dynamics, the other on solid state physics); and $30,000 towards a symposium in Ghana on the state of biology in Africa, being organized under the auspices of the International Council of Scientific Unions. Funding has already been offered for the initial stages of each of these projects, although whether the totals are reached depends on how many countries make good their pledges to the interim fund. **David Dickson**

Mary Lindley

Mary Lindley, an assistant editor of *Nature* since 1966, died last Thursday, 7 May, after some months of illness. She was the longest-serving member of the present staff of *Nature*, and during her service with the journal had successively been responsible for most of the journal's separable functions. She joined the staff immediately after graduation from the University of Oxford, where she had been trained at Somerville College as a biologist with inclinations towards botany.

Mary Lindley suffered throughout her life from the consequences of restricted growth. Her method of dealing with this handicap was to behave as if it did not exist. In this spirit she learned to type, to drive a motor car, to live on her own and to travel widely. Between 1972 and 1975 she worked in the Washington office of *Nature*, handling the flow of scientific manuscripts through that office. When she died, Mary Lindley was the chairman of the Association for Research into Restricted Growth, an organization whose formation she had helped to stimulate.

As a colleague, one of Mary Lindley's memorable attributes was her sardonic wit. Always economical with words, she could in half a sentence puncture an overblown argument — or make a joke. For much of her time on *Nature*, she was the chief defendant of the view that whole organisms have a place in biological studies.

She appeared to be indefatigable. To avoid the worst of the London traffic, it was until very recently her habit to set off in her tiny motor car before the rush-hour began. She would leave the office late, not only to avoid the evening rush-hour but because (she would insist) there was still work to do. Her greatest achievement was that her dealings with those whom she met in the course of her work were never compromised by her handicap; to them, she was merely a very sharp lady who happened to be small.

As an editor, Mary Lindley was the most flexible of people. In her time, she had been responsible for the Book Reviews and the News and Views sections of the journal. For more than a year before she died, she had coordinated, edited and written much of the output of the Nature–Times News Service, which appears in the London *Times*. A few days before she died, cheered by her release from an iron lung and apparently indifferent to the discomforts of a tracheotomy, she asked for a typewriter to be taken to St Thomas's Hospital so that she could begin to write again.

Very few knew of her mounting frustration in the past year about her deteriorating health. Her courage was, however, evident to all.

0028-0836/81/200103-01$01.00

CORRESPONDENCE

Progress in Crete

Sir — The commentary on the current conditions in higher education and research in Greece by Mr E.M. Pantelouris, which appeared in your issue of 1/8 January, came to our attention with some delay. It is, however, instructive to examine the case of the University of Crete as an example of the roadblocks encountered in the actual implementation of innovative ideas and the measure of success that may be expected. This will also demonstrate, we hope, that some of the conclusions of Mr Pantelouris with regard to the University of Crete are not necessarily true.

The operations of the University of Crete started in the academic year 1977–78 with the Department of Mathematics of the School of Sciences at Heraklion and the School of Humanities at Rethymnon. The Department of Physics was added in 1978–79. During this formative phase, instruction was carried out by temporary staff hired on an annual basis and the university was housed in provisional quarters provided in high school buildings.

In spite of these shortcomings, the progress accomplished during the first three years of operation through the efforts of the Provisional Administrative Committee and the enthusiasm of the temporary staff has been remarkable. Some of the major accomplishments include: (1) The organization of a library which already contains about 100,000 volumes in the humanities and 20,000 volumes in sciences. (2) The construction of provisional housing facilities for the schools of science and medicine which are expected to be ready for the academic year 1981–82. (3) The development of a flexible modern curriculum which has been already used successfully. And (4), the organization of student laboratories and the procurement of basic research equipment for the School of Sciences.

Clearly, however, the consolidation of these achievements and the future progress of the university depends critically, at this stage, on the recruitment of a high quality permanent faculty. The Administrative Committee has pushed from the beginning for a departmental structure, in contrast to the chair system, with a high professor to student ratio and a faculty with three professional ranks. The initial appointments will be made by faculty selection committees nominated by the Administrative Committee and consisting of professors from universities in Greece and abroad. These are new and unprecedented features for the Greek university system and it is not surprising that the legislation required for their implementation was time consuming.

This work is now complete and a call for candidacies for 24 positions in the School of Sciences and 14 positions in the School of Humanities has been publicized (see *Nature* classified columns of 29 January). The faculty selection committees have been nominated and the first appointments are expected to be made by late autumn this year. These, of course, are only the first steps in consolidating the structure of the new university and much persistent and enlightened effort will be required especially by the first members of the faculty. The course of the campaign for reforms in the system of higher education in Greece has demonstrated so far that while a diagnosis of flaws and deficiencies is common knowledge, the planning and implementation of effective steps for lasting changes is an agonizing procedure, replete with frustrations but highly rewarding when successful. We hope that the University of Crete will play a significant role in this by setting an example of a high quality modern institution of higher education and research.

A. Kostikas
(Member of Administrative Committee)
Crete University, Greece

Xeroderma list

Sir — The issues raised in Cairns's article "The origin of human cancers"[1] and subsequent comments[2] highlight the importance of studies of rare human disorders as models for human carcinogenesis. In order to gain quantitative information on such questions as the frequency of cutaneous and internal neoplasms in patients with defective DNA repair, we are establishing a registry of xeroderma pigmentosum patients. We welcome correspondence from physicians and scientists who are aware of American patients with xeroderma pigmentosum:

Xeroderma Pigmentosum National Registry,
c/o Department of Pathology,
Medical Science Building, Room C520,
CMDNJ-New Jersey Medical School,
100 Bergen Street,
Newark, New Jersey 07103, USA

Alan D. Andrews
James L. German III
W. Clark-Lambert
Columbia University, The New York Blood Center, NIH and CMDNJ-The New Jersey Medical Center, USA

1. Cairns, J. *Nature* **289**, 353-357 (1981).
2. Trosko, J. *Nature* **290**, 356 (1981).

Be nice to NATO

Sir — Mr Roger R.C. New has recently argued in these columns against publicizing the NATO Advanced Study Institutes in *Nature* on the grounds that it implies approval for "the increasing military orientation of our society" (*Nature* 23 April, p.624). I take precisely the opposite point of view (neither of us questions the scientific and pedagogical value of these institutes).

Surely lasting peace can only be established on this planet if a framework of trans-national judicial, political and economic institutions can be developed to cope with the problems plaguing all mankind. In my view the construction of such institutions must begin with those nations sharing a common democratic tradition — basically those encompassed by NATO and/or OECD. NATO has always been intended as more than a military alliance; I note Article 9 of the Treaty, establishing the "North Atlantic Council" for political consultation, and the meetings of the NATO parliamentarians.

Expansion and encouragement of multinational intellectual, economic and ultimately political ties among the citizens of the world's democracies must be urgently pursued, and the NATO Advanced Study Institutes play a small but positive role in such developments.

William M. Irvine
Chalmers University of Technology, Onsala Space Observatory, Sweden

Computer algebra

Sir — I was delighted to see your recent article on "Algebra made mechanical" (*Nature* 19 March p.198). These symbolic math systems will surely revolutionize scientific computation. However, there is one minor correction to the article: muMATH was coauthored by Albert D. Rich and myself at The Soft Warehouse. Also, to redirect the deluge of enquiries, I suggest that they be sent directly to the distributor: Microsoft Consumer Products at 400 108th Ave. N.E., Suite 200, Bellevue, Washington 98004, USA.

David R. Stoutemyer
The Soft Warehouse, Hawaii, USA

Confusing script

Sir — In the debate stimulated by Halstead's[1] criticisms of the Natural History Museum's exhibit on human evolution, there are two main issues which only partly overlap. First, what cladism entails, and what various authors think about it; and second whether the museum's exhibit presents a fair picture of human evolution, as it is known to science.

Those who disagree with Halstead about the former would do well not to prejudge his assessment of the latter, without seeing the exhibit. B.A. Wood's[2] observations that some early hominids have been seriously misassigned should be noted. So far the archaeological side to the exhibit has passed without comment, but one particular example, in which archaeology and physical anthropology are brought together, speaks volumes for the general tone of the exhibit.

First, it is suggested, reasonably enough, that neanderthals have characteristics not shared by modern man, and vice versa. Then the script continues

> "But because the neanderthals and modern man share these two important characteristics
> — an average brain size of 1,330 ml
> — burial of the dead
> . . . they can be grouped together in the same species — *Homo sapiens*".

This confused conjunction of biometrical and cultural criteria is so weird that little more need be said. But other oddities abound. We are told that we are apes, without any reference to the pattern of characteristics which distinguish us as hominids. Then, the whole, vital, subject of chronology is almost totally omitted.

It seems unlikely that the exhibit represents "official" doctrine of the museum concerning cladism. It represents a good visual show, in which unfortunately "educationalese" has taken over to the extent that the children and other visitors who much enjoy the exhibit are seriously and unnecessarily misled by gross oversimplifications. This is a great pity, for in many ways meticulous care has been taken in the preparation of the exhibit. The desire to simplify in order to spread knowledge is understandable, but over-simplification at the expense of accuracy is counterproductive, and does little to further learning.

John A.J. Gowlett
Research Laboratory for Archaeology and the History of Art, Oxford, UK

1. Halstead, L.B. *Nature* **288**, 208 (1980).
2. Wood, B.A. *Nature* **289**, 8 (1981).

0028-0836/81/200104-01$01.00

Interferon: decline and stall

from Peter Newmark

THE good news to come out of a recent gathering in Rotterdam* was that, given to cancer patients, interferon does have some reasonably consistent clinical effects; the bad news was that there are the side effects of fever, malaise and weight loss. As the number of cases in trials of interferon increases from the tiny to the small and as the trials become randomized, controlled and double-blind, it is clear that interferon is no wonder drug. It remains, however, the case that the number of responses in most trials of interferon is above the threshold for declaring it useless. The consensus view is that interferon will eventually find a permanent place in the limited armoury of the cancer chemotherapist.

The evaluation of interferon in cancer, or for that matter as an antiviral agent, has been unconventional for two main reasons. The first is that it has been available in such small quantities; the second that it has such a marked species specificity that it has always been considered necessary to produce human interferon for human therapy and it has generally been thought pointless to test human interferon on experimental animals. Hence the first very impure products of the heroic Finnish effort, led by K. Cantell, to purify interferon from white blood cells were tested on minute numbers of patients and far too much was sometimes read into the occasional, dramatic response.

Although the supplies of interferon are increasing exponentially, many groups are still testing and reporting handfuls of cases without adequate controls. Among the best current trials of interferon are those of the American Cancer Society (ACS), all of which make use of the Finnish human leukocyte interferon (a mixture of α-interferons). In summarizing the ACS metastatic breast carcinoma trial, Dr E. Borden (University of Wisconsin) included the results of a parallel trial of the same interferon at the M. D. Anderson Hospital in Houston. Of the combined current total of 43 cases, 12 have shown a partial response — defined as one in which there is at least a fifty per cent decrease in the sum of the diameters of all measurable tumours — and lesser responses were recorded in another 6 cases. Despite rumours to the contrary, Borden said the ACS trial will continue, although variations in the size, route and schedule of dosage may be introduced.

The ACS malignant melanoma trial appears less promising, with only one partial response and five lesser responses out of a current total of 43 cases (S.E. Krown). And in only 4 of 23 ACS cases of multiple myeloma was there registered a partial regression of disease (E.F. Osserman, Columbia University). Sadly this does not match conventional therapy and the ACS has decided that, henceforth, interferon will only be tried in combination with more active agents. For breast carcinoma there is still hope that interferon will turn out to be as good as any other single agent (but not as good as a combination of such agents) but it is too soon to be confident even of that conclusion.

Another relatively large trial of leukocyte interferon is on about 50 cases of osteosarcoma (H. Sander, Karolinska Hospital). This trial, unlike its promising predecessor, is being properly controlled and is not yet complete. Sander's impression was that interferon was proving to be better than conventional chemotherapy. Most other reported trials contained less than 10 patients except for a 26-case trial of fibroblast interferon (β-interferon) against neuroblastoma on which it is too early to pass judgement (D. Niethammer, University of Tubingen).

It is also too early to discover the results of the larger trials of interferon as an antiviral prophylactic agent. In particular, a proper trial of the ability of Finnish interferon to prevent herpes and papova virus infections of renal transplant patients (M.S. Hirsch, Boston) will take at least another year to complete. Meanwhile there is still guarded optimism for the use of interferon in the treatment of hepatitis B and both R. Sundmacher (University of Freiburg) and E. De Koning (Utrecht) reported that trifluorothymidine, in combination with Finnish interferon, was markedly more effective against herpes virus-induced dendritic keratitis of the corneal epithelium than was trifluorothymidine alone.

As the clinical trials progress, so does the purity of the interferon tested. Purity can mean one of two things. The first is the extent to which contaminants have been removed; good preparations are now over eighty per cent pure and have a specific activity of over 10^7 international units per mg protein. But purity can also mean the separation of one subtype of interferon from a mixture. The only way to obtain a single member of the eight or so subtypes of α-interferon is from bacteria containing the relevant gene. Production by such means has advanced to the point where at least two small clinical trials of different subtypes have commenced with two hopes. One is that different subtypes will have distinct and complementary activities, as suggested by some *in vitro* tests (N. Stebbing, Genentech); the other is that some subtypes will be without side effects. In that respect it was the impression of T. Merigan (Stanford University) that subtype α_2 (from Genentech/Hoffmann La Roche) produced the usual side effects in man whereas H. Schellekens (TNO, Netherlands) claimed that the same subtype (from Biogen) did not produce the fever that leukocyte interferon did when given to the rhesus monkeys and yet was just as effective in the prophylaxis of vaccinia virus in that species.

Of great promise for the purification of individual interferons is the growing number of monoclonal antibodies which can be used in affinity chromatography. The original monoclonal (*Nature* **285**; 446, 1980) has now been joined by another one against α-interferon (L. Montagnier, Institut Pasteur) and several laboratories now have monoclonal antibodies against β-interferon (H. Erlich, Cetus Corporation; J. Trapman, Erasmus University; H. Hochkeppel, GBF Braunschweig).

Much attention is being turned to the most enigmatic of interferons, γ-inter-

---

Peter Newmark is Deputy Editor of Nature.

*'The Biology of the Interferon System', Rotterdam, April 21–24, 1981; organized by TNO in collaboration with the Erasmus University; proceedings to be published by Elsevier Biomedical around September 1, 1981.

0028-0836/81/200105-02$01.00

feron. Previously known as Type II or immune interferon, this species is produced by cells of the immune system, particularly T cells. Production can be boosted by a variety of inducers and different laboratories have their favourite inducer. For example, R. Falcoff (Institut Curie) favoured anti-lymphocyte serum, H. Kirchner (*Nature* **291**; 79, 1981) uses phytohaemagglutinin or concanavalin A on continuously cultured T cells, J. Vilcek's group swear by a combination of phytohaemagglutinin and phorbol esters (*Proc. natn. Acad. Sci. U.S.A.* **78**; 1601, 1981), most recently mezerein (Y.K. Yip), which has less tumour promoter activity than the one previously used, and others favour protein A from staphylococci (for example, *Nature* **291**; 77, 1981).

Induced cells are beginning to yield sufficient γ-interferon for its purification (or mRNA for its cloning). The big question is whether or not γ-interferon(s) will turn out to be distinct from other species in structure and activity. Conventional wisdom has been that γ-interferon is distinct but in a review of the literature L. Epstein (University of California, San Francisco) expressed caution in view of the impurities of all preparations so far tested. She finds no convincing evidence that γ-interferon has any unique effects on normal cell function and presented evidence that it induced the same 11 peptides in fibroblasts as did α-interferon. There remains, however, the possibility that γ-interferon is quite distinct; it has, for example, at least twice the molecular weight of other interferons (unless this represents polymerization) and is the only one to be unstable at *p*H 2. Unfortunately it also becomes very unstable on purification.

What, then, can be said of the future of the interferons as therapeutic agents? The consensus view appears to be that they will find a place in the treatment and prevention of viral diseases and in the treatment of cancer. For the latter they offer a unique therapeutic approach because they act not as toxins, directly on the tumour cells, but through the immune system, particularly natural killer cells. It is because of this that there is some optimism that γ-interferon, itself from cells of the immune system, will turn out to be a more potent anti-cancer agent than the other interferons. If the problem of side effects (see, for example, G. Scott *et al. Br. med. J.* **282**; 1345, 1981) can be overcome, much larger doses than in the current typical schedule (10^6 units daily for 30 days and thereafter less frequently) may also improve the efficacy of interferons; already some patients have tolerated repeated doses of 100×10^6 units. Although interferon is not the wonder drug that many had hoped, though few had promised, a reasonable, positive view, as expressed by G. Galasso (NIH), is that we should at least be thankful for the arrival of a new drug which is active against diseases for which there are usually few, if any, therapeutic agents. ☐

Palaeoclimates north and south

from M. James Salinger

Following the discovery that changes in Southern Hemisphere sea surface temperatures preceded changes in the Northern Hemisphere by about 3,000 years (J.D. Hays, X INQUA Congress, Birmingham, 1977), research into Southern Hemisphere palaeoclimates has accelerated rapidly. New data, presented at the first CLIMANZ* conference, allow a comparison between palaeoclimatic trends over a large area of the Southern Hemisphere with those of Europe and North America.

In the Northern Hemisphere the last glacial period commenced with a cooling at about 70,000 BP followed by a long period of lower temperatures. This period of cooler climates, continental ice masses and lower sea level was punctuated by a complex series of interstadials, with maximum ice volume at about 18,000 BP. Warming then occurred with very rapid temperature rises between about 15,000 and 12,000BP. A more gradual rise occurred after 10,000 BP with the warmest period in the Holocene at about 6,000 BP.

Oxygen isotope measurements from the Huon Peninsula in northern New Guinea (P. Aharon, Australian National University (ANU)) show that relatively rapid fluctuations occurred in the average annual temperatures between 53,000 and 37,000 BP, with a maximum cooling of about -6°C below present temperatures during a marine regression at 43,000 BP. At 29,000 BP when sea level was relatively high (-45 metres) temperatures were 2.6 ±0.7°C less than at present. These results can be compared with those from deep-sea cores analysed by CLIMAP which suggest a maximum cooling of about 2°C in the western Pacific. Lake data (J. Bowler, ANU) indicate a hydrological regime which was radically wetter than present around 30,000 BP and earlier.

The New Zealand glacial record (T. Chinn, Ministry of Works, Christchurch) suggests that the maximum advance of the Pleistocene occurred around 23,000 BP and continued until after 21,000 BP, with a depression in annual temperatures, deduced from glacial deposits, of approximately 5°C. Corroborative data came from J. Hays (Lamont–Doherty Geological Observatory); faunal evidence from a deep-sea core at 42°S 80°W indicates a sea surface temperature of 8°C below present at 21,000 BP. Hays speculated that the greatest extent of Antarctic sea ice occurred around this time. These results seem to confirm that Southern Hemisphere changes lead those in the Northern Hemisphere by about 3,000 years, at least for the last glacial maxima.

A little later, around 18,000 BP, sand dunes were active in Australia, suggesting a period of increased windiness (J. Bowler, J. Ash and R. Wasson, ANU). Stronger westerlies displaced to the north are indicated by the distribution of quartz grains off the east coasts of Australia and New Zealand (see J. Thiede *Geology* **7**; 259, 1979). These results, together with CLIMAP sea surface temperature data presented by Hays, imply that, at this time when the globe was extremely cold and Northern Hemisphere ice masses were at their maximum extent, the Southern Oscillation, an important feature of the atmospheric general circulation, was much reduced. This would have other implications, not least being a marked reduction in the frequency of upwelling along the South American coast (the El Nino phenomenon).

Comparison of climatic trends in Europe and North America with Australia and New Zealand

Time period (Before present)	Europe and North America	Australia and New Zealand region
40,000–25,000	Fluctuating but cool temperatures	Fluctuating but cool temperatures
23,000–20,000	Cooling	Maximum coldness
18,000–17,000	Maximum coldness	Cold
15,000–12,000	Rapid warming around 13,000–12,000	Rapid warming at about 14,000
10,000–8,000	Warming	Maximum warmth
Approximately 6,000	Maximum warmth	Slight cooling

M. James Salinger is in the Climatic Research Unit, University of East Anglia.

---

*The CLIMANZ (Climatic Change of Australia and New Zealand) conferences aim to bring together groups of workers trying to reconstruct late Quaternary Australia and New Zealand climates.

0028-0836/81/200106-02$01.00

Between 15,000 and 10,000 BP a dramatic change in climate occurred. Geomorphic and palynological evidence from New Guinea indicates that deglaciation started before 14,500 BP with a rise in temperature amounting to 6°C over the entire period (G. Hope, ANU). Glacial changes in New Zealand were extremely rapid, retreating from near-maximum conditions around 17,500 BP to conditions similar to today by 13,500 BP (T. Chinn). At around 15,000 BP there was no substantial area of lowland forest in New Zealand but 5,000 years later both islands were almost completely forested (M. McGlore, Department of Scientific and Industrial Research, Christchurch). In the sub-Antarctic, Hays provided evidence of sea surface temperatures from a deep-sea core at 42°S 80°W. At around 15,000 BP temperatures were near to the glacial minimum and by 10,000 BP they had attained values near to Holocene maximum levels. His data from 42°S suggest that south of the oceanic Polar Front the sea ice regime changed extremely rapidly. Over 300 years at about 14,000 BP there was a sharp reduction in sea ice cover.

The maximum Holocene warming in the CLIMANZ area occurred at 9,000 BP, adding to the mounting evidence that the warmest, most pluvial periods in the past 15,000 years did not occur simultaneously throughout the globe. The warmest period in Europe is thought to be about 6,000 BP. Pollen from Mt Wilhelm and Mt Carstensz in New Guinea (G. Hope) provides evidence that tree-lines there were higher around 9,000 BP than at present. Data from much of southeastern Australia suggest warmer and wetter conditions.

Stratigraphical and pollen records from a number of sites in the southeastern highlands show highest temperatures and precipitation levels at 9,000 BP (A.P. Kershaw, W. Southern, J. Williams and L. Joyce, Monash University), coeval with high lake levels noted at Lake Frome (J. Bowler). At this locality the pollen analyses of G. Singh (ANU) indicate a summer monsoonal rainfall with annual rainfall of perhaps three times the present total. Between 10,000 and 8,000 BP New Zealand had a more complete forest cover than it had at any other time in the past 80,000 years (M. McGlone). The forests in the North Island grew under wet, mild conditions. Maximum forest cover in southern New Zealand, however, coincided with warmer and drier conditions than present. The deep-sea core at 42°S 80°W indicates a minimum of sea ice at this time. Again Southern Hemisphere temperatures in the CLIMANZ area at least seem to lead those in the Northern Hemisphere by about 3,000 years.

The evidence from moisture indicators show striking similarities with anomaly patterns in Australasia associated with an increased Southern Oscillation (Salinger *Mon. Weath. Rev.* **108**; 1892, 1980; Coughlan *Q. Jl R. met. Soc.* **105**; 707, 1979). The suggestion of a high Southern Oscillation Index around 9,000 BP would imply a stronger Walker Circulation and more frequent cold water upwelling off the South American coast and would have other consequences which are open to independent testing at other locations.

Parallels between the spatial patterns of past climates, and those of today, especially associated with important and well defined features of the general circulation such as the Southern Oscillation, provide a relatively simple way of interpreting local changes in a global context and may help us to understand more fully the lack of synchroneity between Northern and Southern Hemisphere climate changes over the past 50,000 years. ☐

Hooked on microtubules

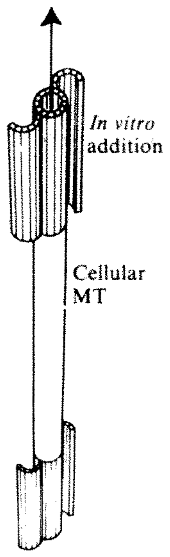
from Jeremy Hyams

In addition to their well documented role in ciliary and flagellar movement, microtubules also participate in many forms of intracellular motility. Phenomena such as the separation of chromosomes at mitosis, the redistribution of pigment granules in fish and amphibian chromatophores and transport along nerve axons are all spatially associated with arrays of microtubules and, in most cases, disrupted by anti-microtubule drugs such as colchicine.

A striking feature of these microtubule-associated movements is their bidirectionality; this property has been taken by many workers to reflect the organization of the underlying microtubules. Briefly, it has been suggested that there exist, in such situations, two families of microtubules with opposed polarities, each capable of exerting force in a given direction. Establishing the polarity of cytoplasmic microtubules has therefore long been regarded as the key to understanding the mechanisms of intracellular motility in much the same way as defining the orientation of the structural proteins of the muscle sarcomere was crucial to defining their interaction during muscular contraction.

Unfortunately for proponents of this view, until now there has existed no unambiguous morphological marker of microtubule polarity analogous to the arrowhead decoration of microfilaments by heavy meromyosin. Experimental dissection of microtubule directionality has therefore progressed only by the less expedient analysis of microtubule elongation *in vitro*. When tubulin subunits are added to a preformed seed (a fragment of microtubule or an organelle such as a flagellar axoneme or basal body), growth is favoured at one end (the fast or plus end) but also occurs to a lesser extent at the other (slow or minus) end (*Nature, News and Views* **284**; 402, 1980).

An important extension of this type of approach was introduced by Heidemann and McIntosh (*Nature* **286**; 517, 1980), who exploited an earlier finding that tubulin, polymerized *in vitro* in a high-ionic strength buffer containing dimethyl sulphoxide, formed not microtubules but sheets of protofilaments (Burton & Himes *J. Cell Biol.* **77**; 120, 1978; Mandelkow & Mandelkow *J. molec. Biol.* **129**; 135, 1979). When Heidemann and McIntosh used existing microtubules to seed this assembly, the polymer formed consisted of whole microtubules to which a C-shaped ribbon of protofilaments had attached, forming a conspicuous hooked appendage. Cross-sectioned images of microtubules grown on *Tetrahymena* basal bodies (see figure), viewed from their plus towards their minus ends, revealed that the hooks had a predominantly right-hand curvature, irrespective of whether microtubules were growing outwards or inwards from the basal body. The method therefore reveals not the direction of growth of these microtubules but, rather, their intrinsic polarity.



In vitro growth of hook-bearing microtubules onto existing seeds. Viewed from the fast-growing or plus end (arrow) the hooks are predominantly right handed irrespective of the direction of growth.

This exciting discovery was limited in its

Jeremy Hyams is in the Department of Botany and Microbiology, University College London.

0028-0836/81/200107-02$01.00

applications to those microtubules which could be elongated *in vitro*. Now, however, by minor modification of their original method, McIntosh and his colleagues have shown that hooks can form directly on existing microtubules (Euteneuer and McIntosh *J. Cell Biol.* **87**; 509, 1980; *Proc. natn. Acad. Sci. U.S.A.* **78**; 372, 1981). In other words, the long-awaited morphological marker of microtubule polarity is at last available.

Application of the new technique to motility-related microtubules has already achieved some notable successes. Decoration of spindle microtubules from both plant and animal cells has sufficiently resolved earlier uncertainties concerning spindle organization to signal the demise of at least one hitherto highly regarded model of mitosis.

Results with other cellular systems have equally demanded that earlier views of microtubules in relation to intracellular motility be reconsidered. In the fish chromatophore, hook curvature was found to be the same for all microtubules and consistent with their plus ends being distal to the cell centre. Thus, even in situations in which transport is clearly bidirectional, the underlying microtubules all possess a single polarity. McIntosh and his colleagues conclude that microtubules do not contribute directly to motility in this system but, rather, form a vectorial framework upon which some other force-generating component is organized. Defining the interaction of microtubules with obvious candidates for this role, such as microfilaments and microtrabeculae, clearly now becomes a major challenge. ☐

Magmatic snap, crackle and pop

from D. H. Tarling

GRAVITATIONAL segregation of crystals out of a magmatic melt has long been considered to be the prime mechanism by which the composition of one or possibly two primary magmas can evolve to account for the observed varieties of igneous rock. This mechanism was originally proposed by Bowen (*Am. J. Sci.* **29**; 175, 1915) as a result of experiments in which olivine glass melts were held for 15 minutes at 1,430°C, before quench cooling, in a pot 2.5 cm high and 1.5 cm wide. The fact that gravitational segregation of individual crystals in a magmatic melt is physically impossible has, however, only been appreciated by a few researchers and the mechanism has only received wide acceptance because there appeared to be no satisfactory alternative.

The actual mechanism by which the fractionation takes place is absolutely vital to any interpretation of the isotopic ratios in igneous rocks. In particular, it is widely accepted that ^{86}Sr/^{87}Sr ratios are low for rocks of mantle derivation. Such a view is based on the assumption that normal oceanic ridge basalts are representative samples of the mantle — that the isotopes themselves have not been partitioned to any significant extent during the formation of the oceanic crust. A recent paper on convective fractionation by Rice (*J. geophys. Res.* **86**; 405, 1981) is therefore critical not only to the understanding of igneous petrological evolution models, but also to the whole basis of geochemical analyses of igneous rock types. It also provides a satisfactory explanation for the occurrence of violent, explosive volcanism associated with the sudden mixing of radically different magma types.

The fundamental objection to crystal segregation is that crystals forming within a cooling magma would take an infinite time to reach the bottom of the magma chamber and so could not accumulate as a crystal mush at the bottom. In any case, the evidence from layered igneous complexes, such as the Skaergaard intrusion of Eastern Greenland, is that the layering takes place on the roof and sides of the chamber as well as on its floor, and it can be repetitive throughout the sequence. To account for such observations, both Hess (*Geol. Soc. Am. Mem.* **80**; 230, 1960) and Wager and Brown (*Layered Igneous Rocks*, Oliver and Boyd, Edinburgh, 1968) proposed that the layering was associated with turbidity currents within the magma chamber, from which individual crystals then segregated. However, although some sedimentary features can be found in some layered complexes, many such features are absent — apart from the intrinsic impossibility of sinking individual crystals through such a viscous medium and the necessity to plate both the walls and ceiling, as well as the floor.

What, then, does Rice propose to replace this widely accepted but impossible model? Essentially he draws on his industrial background in glass and metallurgical processing, combines this with studies by Hurle (*J. Crystal Growth* **13/14**; 39, 1972), and proposes a convective process of fractionation that also explains layering on a range of scales, and shows how magmas of different kinds can be maintained separately and then combined catastrophically to provide explosive volcanism, such as was recently observed in Mount St Helens, and historically at Krakatoa (AD1883) and Santorini (BC1450).

Assuming reasonable physical parameters, the Rayleigh number at which convection would occur within a magma would be obtained in quite small chambers, about 50 m deep, although differences in, for example, viscosity and composition mean that convection could occur on the range of macro- and microscales commonly encountered in geological situations. As the magma convects to lose its heat, there is magma flow past the cooling front at which crystallization is being initiated. Solidification therefore takes place at this front as a function of its temperature and the compositions of the crystallized minerals and the liquid melt. Because the convection is unsteady, there are repeated changes in the composition of the minerals being deposited along the front giving rise to layering. Rice's model predicts that a basaltic magma chamber some 1 km deep would develop layers, each of which would be 5–10 m thick — as observed in the Skaergaard intrusion. One major difference to the previous models is that the layering would occur equally on the floor, ceiling and walls of the magma chamber, as is observed.

Rice's mechanism also explains the apparent close relationship between explosive silicic volcanicity and quieter basaltic eruptions. It has always seemed strange that basaltic lavas can be extruded through the flank fissures of a volcano that was, almost simultaneously, exhibiting violent silicic eruptions at its central cone. (The mechanism may also help to explain the puzzling 'Daly gap' in the range of compositions, between tholeitic and rhyolitic, observed in many oceanic volcanic systems.) In the later stages of the proposed model, there is a high-viscosity, low-density, low-melting point rhyolitic magma overlying a low-viscosity, higher-density, higher-melting point basaltic magma. Rice shows that such a system is stable, apart from some convective intermixing in the boundary zone, until the top of the basaltic magma cools to its crystallization temperature. At this time, the basaltic magma will begin to exsolve volatiles that further reduce its viscosity and its density. This essentially destroys the stability of the system so that there is a near-instantaneous, massive, explosive 'roll-over' of the two magmas, often associated with the generation of super-heated 'flashing', particularly if the magmas are hydrous.

Such a model, if valid, should allow better evaluation of the gestation time of the violent igneous eruptions and clarify how to estimate the potential hazard from the compositions of previously erupted phases. However, one major aspect in this proposed mechanism is that eddy-sized

---

D. H. Tarling is Reader in Palaeomagnetism at the Department of Geophysics and Planetary Physics, Newcastle University, UK.

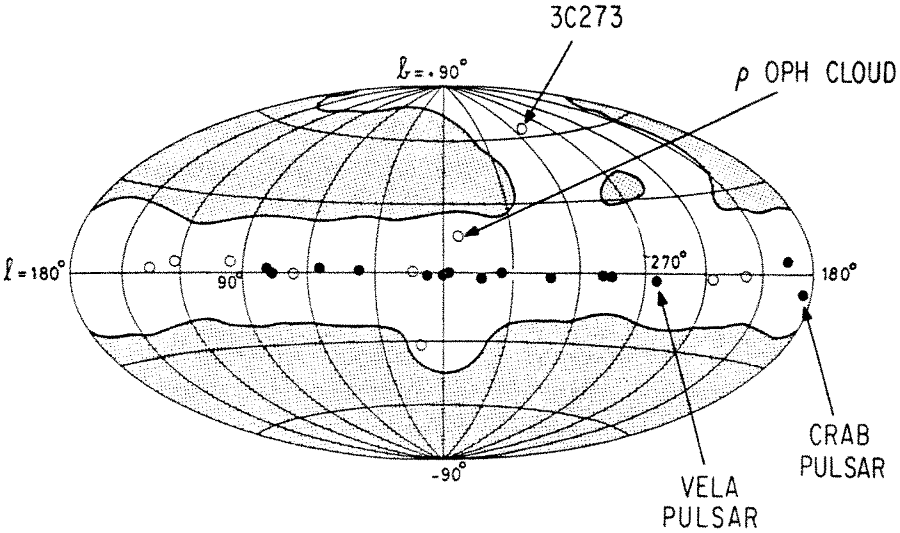
0028-0836/81/200108-02$01.00

A catalogue of unidentified objects

from David J. Thompson

THE sky seen with a gamma-ray telescope offers a unique view of high-energy astrophysical processes. This uniqueness is emphasized by the second COS B catalogue of high-energy gamma-ray sources (Swanenburg *et al. Astrophys. J. Lett.* **243**; L69, 1981) which documents the remarkable lack of obvious correlation between most of the gamma-ray sources and previously known astronomical objects.

The high-energy gamma-ray telescope aboard ESA's COS B satellite has been in operation for nearly six years and data used in compiling the second COS B catalogue span more than three years. The 25 sources of gamma rays with energies above 100 MeV which make up the catalogue may be either stellar or extended sources (with diameters up to 2°). Although the COS B survey does not cover the entire sky, it does include the full galactic plane, along which most of the sources are clustered. There is also a concentration towards the inner part of the Galaxy. In addition to the simple observation that these sources are galactic, their spatial distribution suggests that they lie within 2 to 7 kiloparseconds of the Solar System and, hence, are radiating on the order of 10^{36} erg s^{-1} in the form of high-energy gamma rays.



The second COS B catalogue of high-energy gamma-ray sources, shown in galactic coordinates. Observations covered the unshaded region. Filled circles denote sources with measured fluxes $\geq 1.3 \times 10^{-6}$ photons (>100 MeV) cm^{-2} s^{-1}, while open circles denote sources below this threshold (from Swanenburg *et al.* 1981).

Despite efforts by a number of investigators, only four of the 25 sources have firm identifications. The Crab and Vela pulsars are recognizable by their time signatures, having the same periods as the radio pulsars. The quasar 3C273 and the interstellar cloud complex near the star ϱ Ophiuchus are identified by positional coincidence, which is meaningful only for these high-latitude sources. An extended source, not appearing in the catalogue, has also been identified by COS B with the molecular clouds in Orion (Caraveo *et al. Astr. Astrophys.* **91**; L3, 1980). The remaining 21 members of the catalogue are distinguished by their lack of counterparts at other wavelengths. In particular, the gamma-ray luminosities of these sources are at least an order of magnitude greater than their X-ray luminosities and several orders of magnitude greater than their radio luminosities. This preferential production of gamma rays indicates an unusual astrophysical setting, dominated by large energy transfers. The fact that the identified gamma-ray sources are not all of the same type suggests that the unidentified sources may themselves not be a homogeneous group. The variety of energy spectra seen for the sources in the COS B catalogue is a further indication that two or more classes of gamma-ray source may emerge from future studies. The intriguing possibility is that at least some of these presently unidentified sources may turn out to be a totally new class of galactic object.

Further results can be expected from COS B, particularly in the area of time variability studies, but positive identification of these mysterious sources depends on future observations. The Soviet–French GAMMA-I satellite (to be launched in the next two years) should give improved position information for some of these. The instruments selected for the Gamma Ray Observatory (now scheduled for launch in the latter part of the 1980s) have the broad energy range, good angular resolution and high sensitivity to provide the next breakthrough in gamma-ray source studies. The importance of the COS B catalogue is that it establishes a sizeable number of gamma-ray sources with astrophysically unique properties.

David J. Thompson is in the Gamma Ray and Nuclear Emulsions Branch of the NASA Goddard Space Flight Center, Maryland.

---

convective systems, particularly those associated with magmatic fluids within cracks, are likely to be characterized by a Soret type of isotopic fractionation (as used to produce enriched uranium) that could drastically modify the ratios of even relatively stable, immobile isotopes. The lack of evidence for significant magma chambers beneath the Atlantic and Indian Ocean ridges certainly indicates that the fractionation of the mantle melt to form the oceanic crust probably operated predominantly within tensional cracks in the lithosphere, rather than in a clearly defined magma chamber. The magma chambers thought to have been identified along the East Pacific Rise could, in fact, be areas of the mantle with intensive fractures filled with magmatic fluid, rather than the conventional concept of a large volume of magmatic liquid. Even if magma chambers do exist in such situations, Rice's model could mean that there is much more isotopic fractionation of mantle materials than has generally been recognized. Thus mantle homogeneities postulated on the basis of variation in elements and their isotopes from mid-oceanic ridges to off-centred volcanoes could merely indicate slight differences in the Soret fractionation system in the different areas. Much of the evidence for or against mantle sources for many magmas could similarly be at least gross oversimplifications.

The implications of Rice's model are far-reaching. His paper highlights our ignorance of the fractionation processes in magma chambers, and demonstrates the inadequacy of widely accepted, but incorrect mechanisms for fractionation. It could also mean that we are only just beginning to understand the ways in which igneous rocks are formed and that all previous geochemical interpretations must be re-evaluated. □

0028-0836/81/200109-01$01.00

What are mast cells for?

from T. M. Dexter, R. W. Stoddart and S. T. A. Quazzaz

FROM the time that Paul Ehrlich first described the presence of water-soluble, metachromatically staining granules in certain cells after treatment with Toluidine blue[1] to the present day when Astra blue is used to distinguish the same granules[2], staining properties have been the basis for distinguishing tissue mast cells and basophils from other cell types. Indeed for many years these cells were known only as morphological entities whose numbers were increased in certain pathological conditions. Mast cells are present in all mammals so far studied and have a diffuse distribution, being detected in the gut mucosa, peritoneal cavity, subcutaneous tissue lymph nodes and the thymic capsule. Although the precise relationship between circulating basophils and tissue mast cells is unclear, one may be derived from the other.

The origin of these cells has long been a matter of speculation. Burnett suggested[3] that at least some mast cells represent post-mitotic derivatives of T lymphocytes, citing evidence such as the response of mast cells to T-cell mitogens, mast cell reactions during the response to parasitic infections, and the ease with which mast cells can be generated from cultures of lymphoid tissue. As early as 1963, Ginsburg and Sachs[4] showed that mast cells were regularly produced when thymus cells were cultured on fibroblast feeder layers and others[5,6] suggested that mast cells in the gut mucosa are derived from T cells. However, these suggestions are not compatible with the finding that normal numbers of mast cells are present in the subcutaneous tissue of athymic (*nu/nu*) mice which are devoid of T cells[7].

T. M. Dexter is in the Department of Experimental Haematology, Christie Hospital and Holt Radium Institute, Manchester, and R. W. Stoddart and S. T. A. Quazzaz are in the Department of Experimental Pathology, University of Manchester, UK.

Kitamura and co-workers[8] have used the elegant system provided by the bg^J/bg^J mice, where the peripheral blood leukocytes have highly characteristic giant granules (Chediak–Higashi syndrome). Using the granules as cell markers in marrow transplant chimaeras, they showed that mast cells were derived from marrow precursors, and that mice with a genetically determined macrocytic anaemia (W/W^V mice), associated with a defect in the pluripotential haematopoietic stem cell, also lack tissue mast cells[9]. W/W^V mice are useful in that they will accept marrow grafts–which will cure the anaemia–without the need for prior irradiation to ablate the endogenous haematopoietic system. The availability of these two mouse strains with genetic markers has now paved the way for a direct analysis of mast cell development.

In the experiments reported in this issue of *Nature* (p.159), Kitamura *et al.* have exploited the ability of murine haematopoietic stem cells to undergo clonal proliferation and differentiation and form colonies in the spleens of potentially lethally irradiated mice. A mixture of bg^J/bg^J and wild-type stem cells was injected into wild-type mice. The origin of the resulting spleen colonies was determined using the bg^J/bg^J giant granules as a cell marker, and cells from isolated, individual spleen colonies (containing a 'clonally' pure cell population of either bg^J/bg^J or wild-type cells) were injected into W/W^V mice. Several months later, the blood haemoglobin type was used to determine whether the W/W^V recipients had been reconstituted with cells derived from individual spleen colonies. The animals were then killed and the tissue mast cells and granulocytes examined. In many animals, mast cells which were wholly bg^J/bg^J or wild type appeared in the tissue — and in one mouse the mast cells, erythroid cells and some granulocytes were shown to be bg^J/bg^J derived. These data show fairly conclusively that mast cells and myeloid cells have a common origin in the pluripotent haematopoietic stem cells, but note that the long time interval (15 weeks) between reconstitution of the W/W^V mice and analysis of the mast cells leaves open the possibility that thymus processing is an obligatory step for mast cell development.

This possibility has now been excluded, however, by the recent work of Schrader *et al.*[10]. These authors describe a persisting (P) cell (a name which one can only hope will be replaced before it becomes established!) generated in agar cultures of myeloid progenitor cells and which can be grown in liquid cultures for long periods, provided that the cells are supplied with an appropriate stimulating factor (in this case a conditioned medium derived from concanavalin A-stimulated T cells). In the absence of this factor, the cells die. Such cells contain granules which stain with Astra blue, possess histamine and appear to be equivalent to mast cells. They lack lymphoid markers and can be grown from marrow culture of *nu/nu* mice, from

100 years ago

We take the following from the *Colonies and India:* "To say that a train had been stopped by caterpillars would sound like a Yankee yarn, yet such a thing (according to the *Rangitikei Advocate*) actually took place on the local railway a few days ago. In the neighbourhood of Turakina, New Zealand, an army of caterpillars, hundreds of thousands strong, was marching across the line, bound for a new field of oats, when the train came along. Thousands of creeping vermin were crushed by the wheels of the engine, and suddenly the train came to a dead stop. On examination it was found that the wheels of the engine had become so greasy that they kept on revolving without advancing — they could not grip the rails. The guard and the engine-driver procured sand and strewed it on the rails, and the train made a fresh start, but it was found that during the stoppage caterpillars in thousands had crawled all over the engine, and over all the carriages inside and out."

Messrs. Seimens and Halske gave a public trial last week of their new electric railway, which runs between Lichterfelde and the Cadettenhaus, about six miles from Berlin. The trial is stated to have been most successful. It was in a simple tramcar, with an electric battery totally concealed between the wheels, in connection, through the rails it ran on, with the principal battery at the station. The rails are three feet three inches apart, and exactly resemble those of an ordinary railroad, only the gauge behind narrower. The greatest speed obtained on a distance of about one and a half miles was eighteen English miles an hour. Dr Siemens has proved that if necessary a far greater speed could be obtained, but this is not allowed by the German police authorities. It will not be allowed to proceed at more than nine miles per hour. The railway was opened to the public on Monday.

In connection with the subject of "Fascination," Dr. Otto of Schemnitz, Hungary, writes us of a case which came under his notice. In 1859, when the use of firearms was under stringent regulations in Hungary, peasants often killed hares on the Danubian island Creppel in the following way:— Two peasants would drive in a cart over the reaped fields. On spying a hare (say at two to three hundred paces) they proceeded to drive round it some five or six times in succession. One peasant carrying a long stick at length sprang out, at the moment the cart was behind the hare (the cart continuing its course), and coming up to the animal slowly, killed it without difficulty. It was not uncommon to kill thus as many as six or seven hares in one morning.

from *Nature* **24**, 12 and 19 May, 1881

---

1. Ehrlich *Archs Anat. Physiol.* **3**, 166 (1979).
2. Bloom & Kelly *Histochemie* **2**, 48 (1960).
3. Burnett *Cell. Immun.* **30**, 350 (1977).
4. Grinsburg & Sachs *J. natn. Cancer Inst.* **31** (1963).
5. Ishizaki *et al. J. Immun.* **118**, 211 (1977).
6. Guy-Grand *et al. J. exp. Med.* **148**, 1661 (1978).
7. Keller *et al. Experentia* **32**, 171 (1970).
8. Kitamura *et al. Nature* **268**, 442 (1977).
9. Kitamura *et al. Blood* **52**, 447 (1978).
10. Schrader *et al. Immun. Rev.* **53**, 60 (1980); *Proc. natn. Acad. Sci. U.S.A.* **78**, 323 (1981).
11. Askenase *Prog. Allergy* **23**, 199 (1977).
12. Keller *et al. Immunology* **27**, 1039 (1974).
13. Melmon *et al. Science* **177**, 707 (1972).
14. Bourne *et al. Science* **173**, 743 (1971).
15. Ruitenberg & Elgersma *Nature* **264**, 258 (1976).
16. Azizkhan *et al. J. exp. Med.* **152**, 931 (1980).

0028-0836/81/200110-02$01.00

precursor cells which can also produce granulocytes. These cultured mast cells should allow analysis of the variety of responses in which mast cells are thought to be involved. However, tissue mast cells appear to be functionally heterogeneous and the 'P' cells may represent a developmentally restricted sub-population.

What is the function of mast cells? For many years they have been implicated in hypersensitivity reactions[11] and many parasites, such as schistosomes, are known to stimulate mast cell accumulation. After such infections there is a rapid and dramatic increase in the number of mast cells in the lymph nodes (mainly in the cortical and capsular regions) followed eventually by a decline. The initial accumulation is thought to represent migration of mast cells to the lymph nodes; and the 'disappearance' is believed to be due to degranulation rather than exit via the lymph node hilus[12]. Whether the degranulated cells leave the node in some unrecognized form and recirculate is not known. Normally, mast cells are rather more abundant in the medulla and the deeper paracortex than in the peripheral parts of the node; they are strikingly absent from germinal centres. In some pathological conditions a distribution similar to that seen in parasitic infections is found, again suggesting centripetal migration. What the mast cells are doing in the lymphoid organs is open to speculation.

Most cells contain significant amounts of histamine and many cells, including lymphocytes, have histamine receptors[13]. Histamine inhibits lymphocyte-mediated cytolysis by T cells[14], possibly indicating a role for mast cells in modulating T-cell function. In this context it is significant that T cell- deficient *nu/nu* mice, which have normal or increased numbers of mast cells in subcutaneous tissue, nonetheless show a complete failure of intestinal mast cell response to parasitic infections[15]. However, adequate function can be restored with a thymus graft, perhaps indicating that a T-cell helper function is required. Obviously Schrader's 'P' cells may help to unravel this sort of cooperative response.

Mast cells contain within their granules amines, such as serotonin, and can release large amounts of prostaglandins and leukotrienes. They also specifically produce heparin, a highly sulphated glycosaminoglycan. Heparin has been implicated in the sequestration of lipids by tissue cells and stimulates the migration of capillary endothelial cells[16]. The tissue distribution of mast cells suggests that they may play an important part in the production of heparin and local angiogenesis. Since the successful growth of a tumour depends on its blood supply, this is likely to be of profound importance in the determination of patterns of tumour growth. For example, the growth of spherical tumour masses with necrotic centres might be related to a high density of mast cells immediately around, but not within, the tumour mass.

It is apparent that the careful investigation of the occurrence of mast cells in a variety of pathological conditions is likely to reveal new facets in their role in chronic inflammation, response to injury and the initiation of repair. It is already clear that they are much more than pharmacological time bombs or land-mines simply to be tripped by wandering molecules like IgE. ☐

Leishmaniasis and mouse genetics

from F. E. G. Cox

THE paradox that in some forms of leishmaniasis there is a massive antibody response but not protection, while in others there is no antibody response but excellent protection, is gradually being explained by the use of genetically defined strains of mice.

Leishmaniasis is one of the six diseases selected by the WHO for its Special Programme for Research and Training in Tropical Diseases and is the general name given to infection with any one of some five species or 14 subspecies of the protozoan genus *Leishmania*. Leishmaniasis is transmitted by sandflies and the parasites injected into the human host multiply by repeated binary fission within macrophages in various parts of the body. The manifestations of the disease range from simple, slowly healing cutaneous lesions to fatal disseminated infections involving the viscera. In the Old World the cutaneous form is caused by *Leishmania tropica* and the disseminated visceral form by *Leishmania donovani*. In the New World *L. tropica* is replaced by a number of species that affect not only the skin but also underlying tissues.

At present there are no fully effective drugs or methods of vector control for the prevention of leishmaniasis and no vaccine at all. Fortunately, the parasites can infect mice, so recent studies using these animals have provided a wealth of information that may eventually be useful in the control of leishmaniasis.

Resistance to *L. donovani* and *L. tropica* is genetically controlled — some strains of mice are very resistant to infection while others are susceptible[1-3]. In *L. donovani* infections resistance is controlled by a single autosomal gene on chromosome 1, designated *Lsh*. Differences between resistant (dominant) and susceptible (recessive) mice appear before day 50[4]. Among the susceptible mice some exhibit prolonged infections while others experience a dramatic fall in parasite load. This response is controlled by a gene or genes within or close to the H–2 region of chromosome 17 which allows some innately susceptible mice to recover[5].

L. tropica infections, similarly, are under genetic control with some mice, such as the C57BL, being resistant and others, in particular BALB/c mice which experience fatal disseminated infections, being susceptible[6]. Resistance to *L. tropica* is controlled by a single autosomal gene which is neither the *Lsh* gene involved in *L. donovani* infections nor the H–2 complex, although H–2 genes may be involved late in the infection[6]. Immunity to *L. tropica* involves an antibody-mediated immune response of the delayed hypersensitivity type (DTH) in which the main cells concerned are T lymphocytes (T_D). In susceptible BALB/c mice the lack of immunity is due not to a failure to elicit a DTH reaction but to the induction of suppressor T lymphocytes (T_S), cells which impair the DTH reaction[7]. Early in the infection, DTH is detectable but later it is suppressed. The pattern of events has been elegantly investigated by Howard, Hale and Liew[8]. They subjected BALB/c mice to sublethal irradiation at levels known to delete the precursors of T_S cells and found that the infection, instead of progressing to a fatal conclusion, began to resolve at the same time as those in genetically resistant mice.

Immunity to *L. tropica*, therefore, depends on a DTH reaction which in certain mice can be suppressed by T_S cells. What role, if any, does the humoral immune response play? In the experiments using irradiated mice[8], irradiation at levels that ablate antibody production but not T_S cells had little effect on *L. tropica* infections. Antibody levels in both

---

F. E. G. Cox is Professor of Zoology in the University of London King's College.

1. Bradley, D.J. *Clin. exp. Immun.* **30**, 130 (1977).
2. Handman, E., Ceredig, R. & Mitchell, G.F. *Aust. J. exp. Biol. med. Sci.* **57**, 9 (1979).
3. Howard, J.G., Hale, C. & Liew, F.Y. *Nature* **288**, 161 (1980).
4. Bradley, D.J. *et al. Clin. exp. Immun.* **37**, 7 (1979).
5. Balckwell, J., Freeman, !. & Bradley, D. *Nature* **283**, 72 (1980).
6. Howard, J.G., Hale, C.T. & Chan-Liew, W.L. *Parasite Immun.* **2**, 303 (1980).
7. Howard, J.G., Hale, C.T. & Liew, F.Y. *J. exp. Med.* **152**, 594 (1980).
8. Howard, J.G., Hale, C.T. & Liew, F.Y. *J. exp. Med.* **153**, 557 (1981).
9. Biozzi, G. *et al. Curr. Topics Microbiol. Immun.* **85**, 31 (1979).
10. Hale, C. & Howard, J.G. *Parasite Immun.* **3**, 45 (1981).
11. Perez, H., Labrador, F.T. & Torrealba, J.W. *Int. J. Parasit.* **9**, 27 (1979).

0028-0836/81/200111-02$01.00

irradiated-healing and non-healing mice were similar. A more direct approach to the role of antibody has involved the use of Biozzi high- and low-responder mice. In Biozzi high responders (Ab/H) antibody levels are high and macrophage activity low while in low responders (Ab/L) the reverse is the case[9]. In Ab/H mice infected with *L. tropica,* large, progressive, ulcerating lesions are produced in the presence of high specific antibody titres while in Ab/L mice the lesions are small and healing and antibody levels are minimal[10]. These results confirm unequivocally the widely held belief that in *L. tropica* infections the main protective mechanism is a DTH response and that antibody plays little or no part in recovery. The actual outcome of the infection depends on the interplay of T_p and T_S cells and in due course this information should give us some insight into why some human patients easily overcome the infection and experience only an insignificant lesion while others suffer from spreading and sometimes fatal infections. These findings are relevant not only to *L. tropica* but also to other species. There are indications that similar mechanisms may be involved in infections with the New World species, *Leishmania mexicana*[11], and it may be that the failure to overcome *L. donovani* infections also involves a defect in the DTH mechanism. □

Interstellar grains: models and measurements

from David A. Williams

INTERSTELLAR GRAINS are small solid particles mixed with the interstellar gas which reveal their presence to astronomers in various ways: by causing extinction and polarisation of starlight, or by emitting in the infrared (IR). Extinction is caused by absorption, which heats the grains, or by scattering, which gives rise to an observable, diffuse radiation. Grains also have chemical effects on the interstellar gas, being responsible for conversion of atomic to molecular hydrogen by surface catalysis, and for heavy element depletion. Although grains have been discussed for more than forty years, a recent conference* showed that the questions astrophysicists ask about them have not greatly changed: what are they made of, how are they formed and destroyed, and what is their interaction with gas and radiation?

Astrophysicists have traditionally relied on the curve showing variation of interstellar extinction with wavelength as a pointer to grain identification. In reviewing visual and UV observations Wolstencroft (Royal Observatory, Edinburgh) emphasised the remarkable constancy of the extinction curve in different regions of the Galaxy, especially in visual wavelengths. Some variation does occur in the UV suggesting that (at least) the grain size distribution may vary. Whittet (Preston Polytechnic) showed from a study of the ϱ Oph cloud that, contrary to previous reports, such variations are not accompanied by significant changes in the gas to dust ratio. Measurements of the diffuse (scattered) light in the visual and UV place conflicting constraints on grain properties. Duley (York University, Toronto) and Williams (UMIST) offered a solution in terms of an additional UV component arising from H_2 fluorescence that is not normally allowed for.

Participants were excited to hear from Scarrott (University of Durham) of several peculiar emission features longwards of 5,700Å from the mysterious Red Rectangle. The origin of these features caused much speculation and Greenberg (Rijksuniversiteit te Leiden) suggested an answer based on the laboratory simulation of grain mantles. Ices of simple gases (such as water, methane and ammonia) in approximately cosmic proportions are irradiated by UV and dramatic changes in the material are observed. Flashes of light may be emitted as warm-up occurs and the residue has optical features which Greenberg proposes are the diffuse bands.

Unidentified spectral features associated with grains also abound in the IR. Whittet reviewed recent observations of emission features from HII regions, planetary nebulae, galactic nuclei, and the absorption features in the cool interstellar gas. Most participants were convinced that the prominent 9.7 μm feature (seen in emission at high temperature and in absorption) is to be attributed to the SiO bond in refractory material such as amorphous silicates or oxides. The 3.1 μm feature, seen only in absorption in dark clouds, was generally attributed to molecular mantles accreted by cold grains.

A number of other IR absorption and emission features in the range 3–12 μm are known and two alternative interpretations were given. Allamandola (Rijksuniversiteit te Leiden) presented results of laboratory measurements of absorption features from ices such as H_2O, CO and CH_3OH, and identified them with observed absorption features at 3.4, 3.5, 6.0 and 6.8 μm. Duley and Williams suggested that radicals bound at reactive sites on amorphous carbon grains produce these and other absorption features, and also emission features from warm grains. The precise spectral properties of such grains will change according to their environment. Clearly, the flood of IR observations is providing a set of tight constraints on grain composition.

These and other constraints were discussed by Greenberg, who suggested that all observations so far were consistent with models of amorphous silicates or oxides, carbons, dirty ice mantles and photoprocessed material. The interaction of gas and grains can be studied in detail when the grain material is specified, for then the vast surface chemical literature can be tapped. Duley showed how effectively this can be done. He described a consistent model, including depletion, chemical reactions and optical properties, with the attractive feature that the chemical and optical nature of grains change as grains move from one environment to another. Apparent changes in grain properties may merely reflect their physical state.

The formation of grains is indicated astronomically in a number of sources: late-type giants and supergiants, other special stars, planetary nebulae, novae and supernovae. Some problems remain in understanding nucleation processes. Mechanisms of destruction (sputtering, evaporation and shattering in collision) seem able to account for the life times of grains. Evans (University of Keele) reviewed observations of a nova exhibiting grain formation. There are some difficulties in explaining why the grain temperatures increased after the nova had already started to decline in brightness. There are also some quite puzzling observations of ejection in WR stars in which optical and IR intensities decline with time but the measurement at 3.4 μm remains high.

The Durham group, in work described by Smith, demonstrated that sophisticated modelling of optical polarisation in reflection nebulae is particularly rewarding in understanding the nebulae and the grains. Robinson (Queen Mary College, London) convinced the workshop that in such modelling, especially of IR sources, workers must account correctly for radiative transfer or quite spurious interpretation concerning grains and source properties may result.

The meeting ended in a mood of some optimism. Although much controversy remained it was felt that new observations, particularly in the IR and UV, allied to sophisticated modelling, will enable significant progress to be made in the next few years. □

David A. Williams is in the Department of Mathematics, University of Manchester Institute of Science and Technology.

---

*A workshop on interstellar grains was held in Manchester on April 7–8, 1981.

0028-0836/81/200112-01$01.00

ARTICLES

Episodic granite plutonism in the Scottish Caledonides

Otto van Breemen

Scottish Universities Research and Reactor Centre, East Kilbride, Glasgow G75 0QU, UK

Brian J. Bluck

Department of Geology, The University, Glasgow G12 8QQ, UK

Boulders from a Lower Devonian conglomerate complete the record of a phase of Ordovician plutonism associated with island arcs and blue schist metamorphism, indicating high angle subduction. Following a phase of sedimentary accretion, late Silurian plutonism was the result of low angle subduction. The two granite suites have distinct isotopic trends.

BOULDERS of two-mica granite which occur as clasts in the Lower Old Red Sandstone conglomerates along the Highland Boundary Fault (Fig. 1) have yielded Rb–Sr muscovite ages of ~460 M yr confirming the existence of a widespread suite of subduction-related Ordovician granites south of the Great Glen Fault. A voluminous suite of late Silurian to early Devonian granites can be related to the subduction of comparatively younger, thinner, hotter and more buoyant oceanic crust. The intervening hiatus in plutonism coincides with the peak of subduction accretion in the Southern Uplands, further demonstrating the necessity of subducting wet sediments to generate granitic magmas. For the Ordovician suite, the sedimentary record completes an isotopic trend of increasing $^{87}Sr/^{86}Sr$ ratios away from the trench which can be related to the increasing involvement of upper crust. The later granites show a trend towards more unradiogenic lead away from the trench which can be attributed to the increasing involvement of lower crust.

The Lower Old Red Sandstone rocks of the Midland Valley of Scotland comprise a thick sequence of late orogenic alluvium and intercalated volcanic rocks[1]. The sequence at Stonehaven (Fig. 1*a*), possibly exceeding 8 km in thickness, contains abundant conglomerates which are dominated by igneous clasts[2] (Fig. 1*b*). Andesitic and basaltic clasts are particularly abundant and often less well rounded than accompanying granite clasts of similar size. Granite which locally makes up ~25% of the clasts, ranges from microgranite to pegmatites and is invariably primary muscovite bearing. These clasts have generally been considered to have been derived from the late Silurian–early

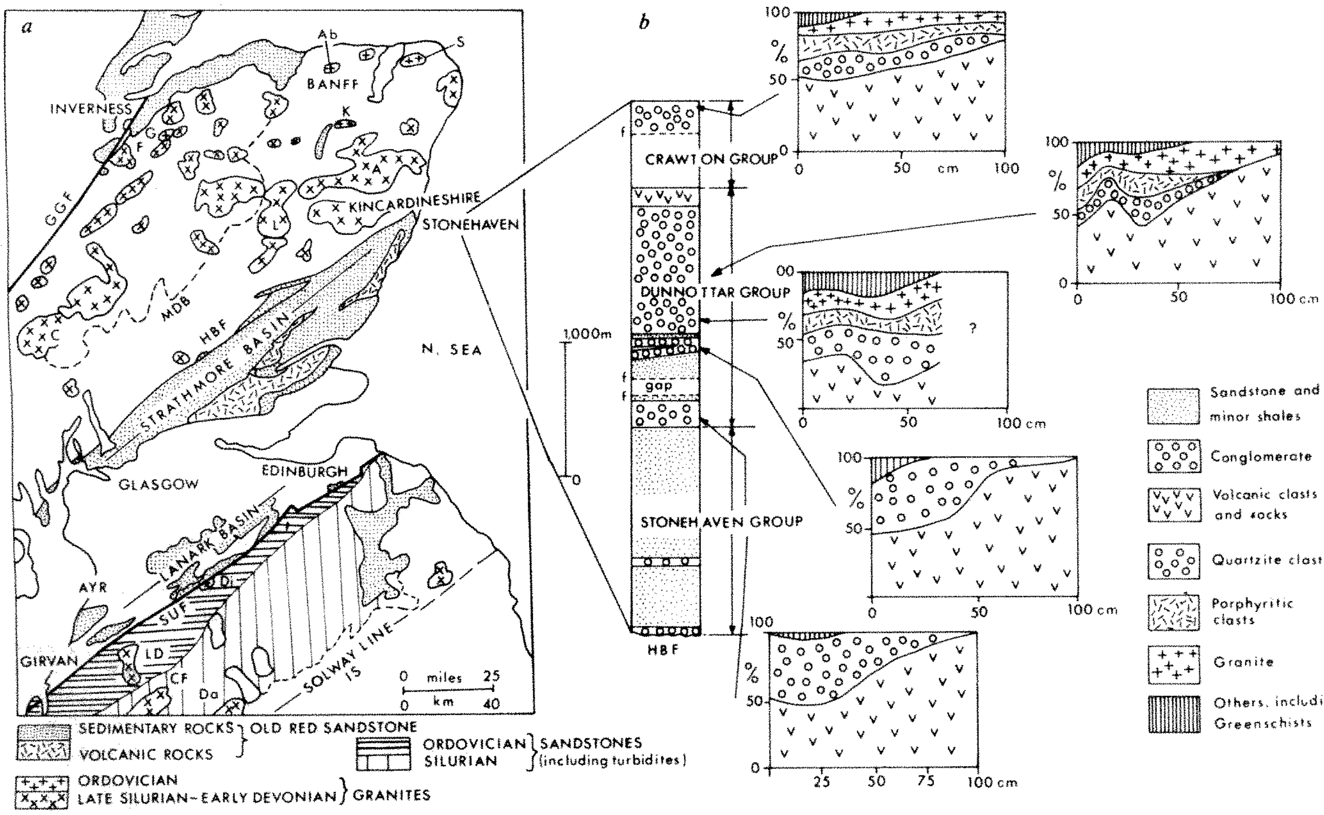


Fig. 1 *a*, Showing distribution of Caledonian granites in Scotland, the major tectonic lines and localities referred to in the text. Ordovician granites: Ab, Aberchirder; K, Kennethmont; S, Strichen; G, Glen Kyllachy. Silurian–Devonian granites: Da, Dalbeattie; CF, Cairnsmore of Fleet; LD, Loch Doon; Di, Dinstinkhorn; A, Aberdeen–Hill of Fare; L. Lochnagar; C, Cruchan; F, Foyers. Tectonic boundaries: IS, Iapetus Suture; SUF, Southern Uplands Fault; HBF Highland Boundary Fault; MDB, Moine Dalradian Boundary; GGF, Great Glen Fault. *b*, Stratigraphic sequence at Stonehaven–Crawton[2] and distribution of rock types-by-size through the sequence.

0028-0836/81/200113—05$01.00

Table 1 Rb–Sr isotopic data

Sample	Rb (p.p.m.)	Sr (p.p.m.)	^{87}Rb/^{86}Sr (atomic)	^{87}Sr/^{86}Sr$_i$ (atomic)	Mineral age (Myr) λ ^{87}Rb = 1.42×10^{-11} yr^{-1}	(^{87}Sr/^{86}Sr)$_i$ (atomic)
R.C. 1913						
Whole-rock	133	48.4	7.999	0.77254 ± 4		0.7196 ± 6
Muscovite	529	4.23	474.1	3.8585	465 ± 5	
Biotite	867	19.7	138.2	1.5536	421 ± 5	
R.C. 1916						
Whole-rock	172	116	4.287	0.74295 ± 6		0.7150 ± 3
Muscovite	521	22.5	70.15	1.1721	459 ± 5	
R.C. 1917						
Whole-rock	110	217	1.466	0.72127 ± 7		0.7116 ± 2
Muscovite	344	27.1	37.61	0.95881	461 ± 5	
Biotite	481	37.0	38.57	0.96406	459 ± 5	
Biotite	491	35.0	41.08	0.97964	458 ± 5	
R.C. 1919						
Whole-rock	142	206	1.999	0.72326 ± 4		0.7103 ± 2
Muscovite	339	20.8	48.68	1.0270	457 ± 5	
Biotite	649	17.8	112.7	1.4236	444 ± 5	
R.C. 1920						
Whole-rock	113	83.1	3.931	0.73790 ± 6		0.7122 ± 3
Muscovite	382	9.74	122.4	1.5127	459 ± 5	

Devonian granites to the north[3,4]. However, the muscovite-bearing granite clasts become less abundant upwards in the succession.

Isotope analyses

Rb–Sr isotope dilution analyses (Table 1) follow analytical methods described elsewhere[5]. Rb–Sr muscovite ages for the four muscovite biotite granites RC 1916, 17, 19, 20 are concordant at 459 ± 5 Myr. The Rb–Sr muscovite age for pegmatitic sample RC 1913 may be slightly older at 465 ± 5 Myr. Rb–Sr biotite ages scatter down to 421 ± 5 Myr. The biotite age for sample RC 1917 is, however, concordant with the muscovite ages.

The Rb–Sr muscovite ages are here interpreted in terms of the time of emplacement of these underformed granites and pegmatite for two reasons. First, in the central Highlands and within the axial region of the Caledonian orogen, Rb–Sr muscovite systems have survived all phases of Caledonian deformation and associated amphibolite facies metamorphism[6]. Second, the concordance of the biotite and muscovite ages for sample RC 1917 indicates that regional temperatures in at least some of the host rocks could not have greatly exceeded 300°C, the estimated blocking temperature for the Rb–Sr isotopic system in biotite[7]. The younger, scattered biotite ages are believed to have been caused by alteration to chlorite during deposition in lower Devonian time.

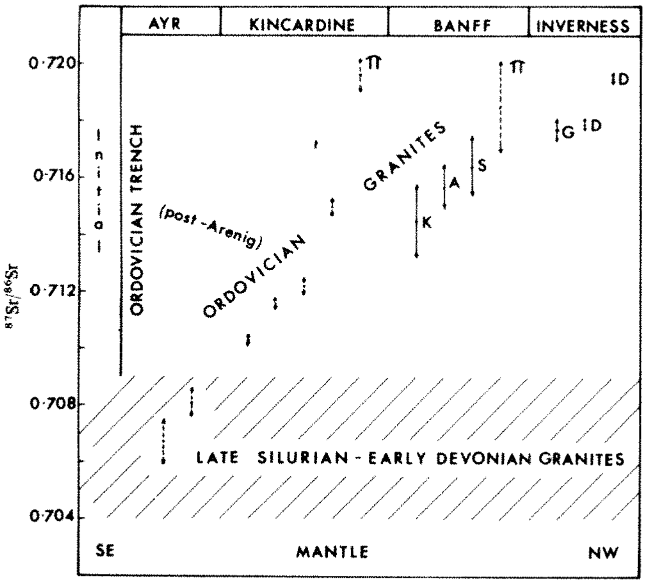
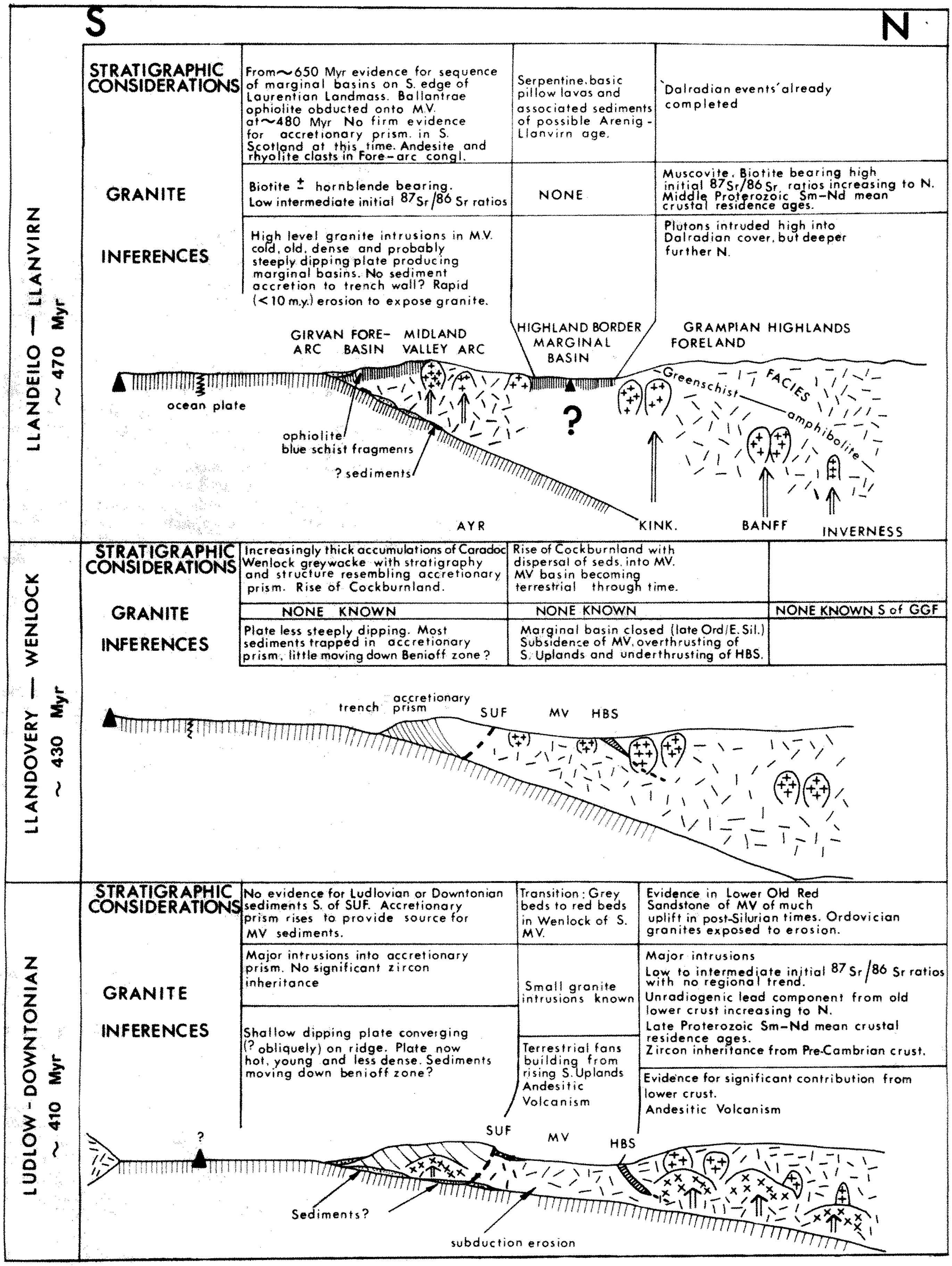


Fig. 2 Cumulative frequency plots of initial ^{87}Sr/^{86}Sr ratios for post-Arenig Ordovician granites in Scotland. From left to right regions have been arranged from the Ordovician trench into the interior of the continent. Dashed lines indicate uncertainties on individual whole rocks: clasts, granodiorite dykes (D) or pegmatities (π). Solid lines indicate uncertainties on whole rock isochrons. See Fig. 1.

The wide range in initial ^{87}Sr/^{86}Sr ratios obtained is interpreted in terms of increasing upper crustal involvement. The chemistry of the rocks bears little systematic relation to the initial ^{87}Sr/^{86}Sr ratios which suggests that the magmas sampled by the boulders did not involve simple mixing between two end members. Concentrations of SiO_2 range from 71 to 74 wt%, Na_2O from 3.1 to 3.9 wt%, K_2O from 3.5 to 4.8 wt% and normative corundum from 0.1 to 2.6 wt%. The wide range in ^{87}Sr/^{86}Sr ratios combined with the lack of evidence for a mixing pattern suggests that the boulders came from more than one intrusion. Thus the tight grouping of Rb–Sr muscovite ages at ~460 Myr supports the evidence from the Stonehaven clast profile (Fig. 1) that these late Ordovician granites represent a distinct pulse of magmatism.

Granite clasts

The provenance of the ~460 Myr granite clasts is not known. The conglomerates in which they occur thicken onto the north-west limb of the Strathmore Syncline suggesting a source to the north-west, a view supported by the presence of Highland-type rocks in the conglomerates. In terms of durability during transportation, schists and volcanic rocks would probably break down quite quickly, granite less quickly and quartzite least of all[8]. On this basis it is estimated that granite formed some 10–15% of the exposed ground in the source at about the time the Dunnottar Group (near Dunnottar Castle) was laid down. The present areas to the north of Strathmore (in southern Aberdeenshire) contain some 43% granite bedrock by area (taking a strip of country 20 km to north-west of the Highland Boundary Fault for 60 km south-west from Stonehaven). These Silurian–Devonian granites have not been found in the Devonian conglomerates. The nearest rocks of ~460 Myr age are found some 60 km to the north in Banff[9]. In view of the size of the granite boulders (up to 80 cm) transport over this distance is unlikely, and a more proximal source either in the Midland Valley or in the southern Highlands is envisaged. Given no substantial lateral movement on the Highland Boundary Fault and accepting the stratigraphical evidence, we suggest that these early plutons were emplaced high into the Dalradian cover to the north, among rocks which were largely in greenschist facies metamorphism. These greenschist rocks provided some of the



LLANDEILO — LLANVIRN ~ 470 Myr

STRATIGRAPHIC CONSIDERATIONS	From ~650 Myr evidence for sequence of marginal basins on S. edge of Laurentian Landmass. Ballantrae ophiolite obducted onto M.V. at ~480 Myr No firm evidence for accretionary prism. in S. Scotland at this time. Andesite and rhyolite clasts in Fore-arc congl.	Serpentine, basic pillow lavas and associated sediments of possible Arenig-Llanvirn age.	'Dalradian events' already completed
GRANITE	Biotite ± hornblende bearing. Low intermediate initial $^{87}Sr/^{86}Sr$ ratios	NONE	Muscovite, Biotite bearing high initial $^{87}Sr/^{86}Sr$ ratios increasing to N. Middle Proterozoic Sm–Nd mean crustal residence ages.
INFERENCES	High level granite intrusions in M.V. cold, old, dense and probably steeply dipping plate producing marginal basins. No sediment accretion to trench wall? Rapid (< 10 m.y.) erosion to expose granite.		Plutons intruded high into Dalradian cover, but deeper further N.

LLANDOVERY — WENLOCK ~ 430 Myr

STRATIGRAPHIC CONSIDERATIONS	Increasingly thick accumulations of Caradoc Wenlock greywacke with stratigraphy and structure resembling accretionary prism. Rise of Cockburnland.	Rise of Cockburnland with dispersal of seds. into MV. MV basin becoming terrestrial through time.	
GRANITE	NONE KNOWN	NONE KNOWN	NONE KNOWN S of GGF
INFERENCES	Plate less steeply dipping. Most sediments trapped in accretionary prism, little moving down Benioff zone?	Marginal basin closed (late Ord/E. Sil.) Subsidence of MV, overthrusting of S. Uplands and underthrusting of HBS.	

LUDLOW - DOWNTONIAN ~ 410 Myr

STRATIGRAPHIC CONSIDERATIONS	No evidence for Ludlovian or Downtonian sediments S. of SUF. Accretionary prism rises to provide source for MV sediments.	Transition: Grey beds to red beds in Wenlock of S. MV.	Evidence in Lower Old Red Sandstone of MV of much uplift in post-Silurian times. Ordovician granites exposed to erosion.
GRANITE	Major intrusions into accretionary prism. No significant zircon inheritance	Small granite intrusions known	Major intrusions Low to intermediate initial $^{87}Sr/^{86}Sr$ ratios with no regional trend. Unradiogenic lead component from old lower crust increasing to N. Late Proterozoic Sm–Nd mean crustal residence ages. Zircon inheritance from Pre-Cambrian crust.
INFERENCES	Shallow dipping plate converging (? obliquely) on ridge. Plate now hot, young and less dense. Sediments moving down benioff zone?	Terrestrial fans building from rising S. Uplands Andesitic Volcanism	Evidence for significant contribution from lower crust. Andesitic Volcanism

Fig. 3 Reconstruction of the Southern Uplands–Grampian Highlands during Ordovician–Silurian time, showing facts and inferences referred to in text. See Fig. 1: MV, Midland Valley; HBS, Highland Border Series.

associated clast materials in the conglomerates. As with other plutonic bodies and as suggested by the clast profile, convergence of the margins downwards would result in a balloon shaped plutonic mass[10], which might almost disappear as the land surface was eroded down.

Granite clasts dated[11] at ~470 Myr and associated with an influx of andesitic fragments[12], both derived from the southern Midland Valley, have been related to the late Ordovician two-mica granites in the Dalradian of Banff and Aberdeenshire. In the latter region ages range from 475±5 Myr for the Stichen Granite (U–Pb on monazite[13]) to 453±4 Myr and 444±9 Myr for the Kennethmont and Aberchirder granites (Rb–Sr whole rock isochrons[9]). A regional suite of tourmaline-bearing pegmatites have been dated at 463±6 Myr (Rb–Sr muscovite isochrons[14,15]). All these intrusions are post-tectonic. Further north-west, in the Moine schists below the Dalradian metasediments, late tectonic pegmatites and a small two-mica granite (the Glen Kyllachy Complex) have been dated at 445±5 Myr and ~443 Myr (Rb–Sr muscovite and whole rock isochron ages[5]). If the Rb–Sr isotope data are taken alone, there seems to be a ~30 Myr, decreasing age trend from the south-east to the north-west away from the trench. The Stichen Granite in north-east Aberdeenshire with a reliable U–Pb monazite age of 475±5 Myr seems to be older. Even so, the age of this pluton falls within the ~35 Myr range of late Ordovician granite emplacement.

Discussion

As the Ordovician granites extend south-east of the Highland Boundary Fault where there is no evidence for the Grampian orogenic event, this suite has been interpreted in terms of being related to subduction[11]. Yet high initial $^{87}Sr/^{86}Sr$ ratios and various geochemical features have been presented as evidence that the two-mica granites of north-east Aberdeenshire and the Glen Kyllachy complex have been derived from metasediments[5,9,16]. The new data from the denuded granites along the Highland Boundary Fault provide evidence for a common origin of all the middle to late Ordovician granites and andesites and a link between subduction and high level melting.

The ~460 Myr granites are clearly younger than those migmatites north of the Highland Boundary Fault which are of ~510 Myr Grampian age[17] or older[18]. In view of their proximity to the Highland Boundary Fault it would be difficult to envision their generation through burial metamorphism of Grampian age[19]. Also these granites slightly predate the Caradocian marginal basin closure along the Highland Boundary Fault (Fig. 3, see ref. 20). Their generation and rise past the present erosion level almost certainly required a subcrustal heat source.

Burnham[21] has shown that mantle-derived basic magma can produce extensive partial melting and assimilation of hydrous rocks of felsic composition. Such a process is analogous to the mechanism of magma ascent and country assimilation as developed by Bowen[22,23]. Thus the rise of a magma continues while a net balance of heat is maintained from that lost by melting and or assimilation of country rocks and gained by the latent heat of crystallization of a more basic residuum. The assimilation crystal fractionation process is not unique to the original magma and as the magma bores its way upwards, the major element chemistry will approach that of a partial melt of the country rocks.

Initial Sr isotopic ratios for all the late Ordovician granites are plotted in Fig. 2. Ratios from each area are presented as cumulative frequency plots for increasing values and the north-west is represented by the right-hand side of the diagram. The granites show a systematic increase in $^{87}Sr/^{86}Sr$ ratios as one passes from the Ordovician trench to the interior of the continent. The age and initial ratio pattern of the late Ordovician granites thus seems to be systematic. Although the effects of ultrametamorphism may have been significant, it is reasonable to conclude that the process which relates all these granites, on both sides of a marginal basin, is subduction. The north-westwards increasing $^{87}Sr/^{86}Sr$ ratios could then be attributed to the increasing thickness of older continental crust above the Benioff zone[24].

Clearly the high temperatures of metamorphic rocks in a post-tectonic setting when geotherms are raised in the early stages of uplift[25] should by the above heat balance mechanism allow the magmas to rise high into the upper crust as well as increase the proportion of melt derived from the upper crust. Also, in view of the westward orogenic diachronism across the Grampian region[5] magmatism may have persisted later in the north-west. In the latter region earlier magmatism may be represented by late tectonic 'discordant migmatite'[26]. Also the apparent westward migration of plutonism is probably partly the result of the time constant of subduction.

Variations in the nature and extent of plutonic and volcanic activity have been related to differences in characteristics of the ocean plate being consumed at the trench[27–29]. In Recent to Jurassic examples, where adjacent ocean lineation patterns allow reconstructions of this type, the relationship between characteristics of ocean plate and plutonic activity is still equivocal[30]. In the Caledonian cycle, changes in the characteristics of the Iapetus plate can only be inferred from the stratigraphical record in southern Scotland. Criteria for the recognition of these changes are given in Fig. 3. Vlaar and Wortel[31] and Molnar and Atwater[32] have shown that a relationship exists between angle of subduction and age of plate–older plates are thicker, colder and denser and therefore dip more steeply than the younger. These authors also suggest that island arcs and accompanied marginal basins are produced by the steeply dipping subduction zones.

The geological record in southern Scotland suggests the presence of volcanic arcs and marginal basins in Arenig to Llandeilo times[33] together with possibly older blue schist, from which it might be inferred that the Iapetus plate was cold and steeply dipping. It was during this interval that the early granites (discussed here) were emplaced. These are high level, relatively low volume granites showing much upper crustal involvement. At this same time, despite evidence for rapid uplift in the Midland Valley source[11], the volume of trench accretion does not seem to have been substantial[34]. Some of this sediment may have been consumed in the subduction zone.

In the interval Caradoc–Ashgill, the beginning of sedimentary accretion coincides with decreasing granite magmatism moving away from the trench. The following Llandovery and Wenlock stages are typified by a major accretionary prism and during this

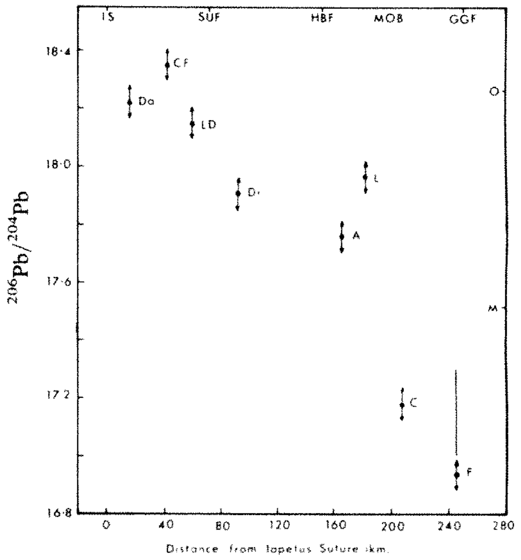


Fig. 4 Plot of age corrected $^{206}Pb/^{204}Pb$ ratios for feldspar separates from Silurian–Devonian granites of Scotland[45] versus distance from the Iapetus Suture. See Fig. 1. Doe and Zartman's[47] mantle and 'orogene' values of $^{206}Pb/^{204}Pb$ 400 Myr ago are represented by M and O. These authors estimate Upper and Lower crustal values at 18.71 and 16.67 respectively. Lewisian rocks yield values in the range 13.41–16.91 (ref. 48).

time, no granite intrusions have been recorded south of the Great Glen Fault. Abundant calc-alkaline plutonism at this time north of the Great Glen Fault can be explained in terms of separate subduction histories before the opposing Highland regions were juxtaposed by sinistral movements along this transcurrent fault[5,35].

The next period of major plutonism south of the Great Glen Fault took place between ~415 and 390 Myr. (Fig. 1)[13,36–39]. There is no record of any sediment accretions at this time in the south of Scotland. Although there may be many reasons for this absence of sediments north of the Iapetus suture, it is possible that they were taken down the subduction zone as in the former period of major plutonic activity. Indeed, the patterns of granite plutonism and sedimentary accretion observed in California[40], Alaska[41] and Scotland support continuous crustal recycling[42,43].

At this stage it is possible that together with the sediments on the ocean floor, part of the accretionary prism was also abraded into the subduction zone. A hypothesis of a mixed derivation of sedimentary and mantle material to make up the voluminous Silurian–Devonian granites is consistent with the Sr (ref. 37) and O isotope data[38,44]. However, common lead isotope data from feldspars show that lower crustal materials were also involved in the formation of these granites[45]. As lower crustal rocks generally have low $^{87}Sr/^{86}Sr$ ratios, the absence of a geographical trend in this parameter (Fig. 2) is not diagnostic. The common lead isotope evidence is supported, however, by more recent Sm–Nd isotope evidence[46].

With the separate history of the later granites on either side of the Great Glen Fault, the Pb isotope trend[45] for the granites to the south can be shown to be surprisingly consistent. On Fig. 4 the $^{206}Pb/^{204}Pb$ ratios of the later granites south-east of the Great Glen Fault have been plotted against distance from the Iapetus suture. More data would be desirable. However, note that four $^{206}Pb/^{204}Pb$ ratios from the Foyers complex (in the range 17.0–17.3) reported by Pankhurst[39] and showing the internal variability in a single pluton are consistent with the trend of northwestwardly decreasing ratios. The consistent downward trend of this ratio away from the Iapetus suture is best interpreted in terms of an increasingly unradiogenic (Pb) lower crustal input into the granites. In view of the relative Pb concentrations of upper and lower crustal rocks, only limited amounts of upper crustal rocks are likely to have been assimilated[47].

The trend of upper crustal (radiogenic Sr) involvement of the Ordovician granites seems to have been repeated with respect to the lower crust (unradiogenic Pb) in the case of the Silurian–Devonian granites. This feature can be partly explained by the large amount of Highland uplift, erosion and cooling which occurred in the intervening (Silurian) time. At the same time, the greater volume of the later granites, which may have swamped the upper crustal contribution, argues for high temperatures in the lower crust and the mantle[49]. The voluminous nature of these granites combined with the abundance of andesitic volcanism, often directly relatable to the plutons, suggests that in the final stages of closure of Iapetus the continent over-rode comparatively young, hot ocean floor at a low angle of subduction. This interpretation is also consistent with the absence of island arcs in the Silurian.

We thank J. Hutchinson for help with the analyses and A. N. Halliday who contributed to discussions on granite volume in relation to angle of subduction. The Isotope Geology Unit of the Scottish Universities Research and Reactor Centre is supported by the NERC.

Received 26 November 1980; accepted 20 February 1981.

1. Bluck, B. J. *Geol. J. Spec. Issue* **10,** 249–278 (1978).
2. Armstrong, M. & Paterson, I. B. *Int. geol. Surv.* Rep. No. 70/12 (1970).
3. Campbell, R. *Rep. Br. Ass. Advmt Sci.* **96,** 554–555 (1929).
4. Waterson, C. D. in *The Geology of Scotland* (ed. Craig, G. Y.) 296–308 (Oliver & Boyd, Edinburgh, 1965).
5. van Breemen, O. & Piasecki, M. A. J. (in preparation).
6. Piasecki, M. A. J. & van Breemen, O. *Nature* **278,** 734–736 (1979).
7. Purdy, J. W. & Jäger, E. *Mem. Instn. Geol. Miner. Univ. Padova* **30,** (1976).
8. Kuenen, P. H. *J. Geol.* **64,** 336–368 (1956).
9. Pankhurst, R. J. *Bulg. geol. Soc. Am.* **85,** 345–350 (1974).
10. Pitcher, W. S. *J. geol. Soc. Lond.* **136,** 627–662 (1979).
11. Longman, C. D., Bluck, B. J. & van Breemen, O. *Nature* **280,** 578–581 (1979).
12. Kelling, G. J. *J. geol. Soc. Lond.* **117,** 37–75 (1961).
13. Pidgeon, R. T. & Aftalion, M. *Geol. J. Spec. Issue.* **10,** 183–220 (1978).
14. van Breemen, O. & Boyd, R. *Scott. J. Geol.* **8,** 115–120 (1972).
15. Lappin, M. P., Crane, A., Aftalion, M. & van Breemen, O. (in preparation).
16. Bursewil, M. T., Pankhurst, R. J. & Wadsworth, W. J. *Miner. Mag.* **40,** 363–376 (1975).
17. Pankhurst, R. J. & Pidgeon, R. T. *Earth Planet. Sci. Lett.* **31,** 55–68 (1976).
18. Sturt, B. A., Ramsay, D. M., Pringle, I. R. & Teggin, D. E. *J. geol. Soc. Lond.* **134,** 41–44 (1977).
19. Harper, C. T. *Scott. J. Geol.* **3,** 46–66 (1967).
20. Beckinsale, R. D. in *Origin of Granite Batholiths, Geochemical Evidence* (eds Atherton, M. P. & Tarney, J.) 34–44 (Shiva, Copenhagen, 1979).
21. Burnham, C. W. in *Geochemistry of Hydrothermal Ore Deposits* (ed. Barnes, H. L.) 71–136 (Wiley, New York, 1979).
22. Bowen, N. L. *The Evolution of Igneous Rocks* (Dover, New York, 1956).
23. Taylor, H. Jr *Earth planet. Sci. Lett.* **47,** 243–254 (1980).
24. Miller, C. F. & Bradfish, L. J. *Geology* **8,** 412–416 (1980).
25. England, P. C. & Richardson, S. W. *J. geol. Soc. Lond.* **136,** 489–495 (1977).
26. Piasecki, M. A. J. *Scott. J. Geol.* **11,** 87–115 (1975).
27. Gilluly, J. *Bull. geol. Soc. Am.* **84,** 499–514 (1973).
28. Dickinson, W. R. *Rev. Geophys. Space Phys.* **8,** 813–860 (1970).
29. De Long, S. E. & Fox, P. J. in *Island arcs, Deep Sea Trenches and Back-Arc Basins* (eds. Talwani, M. & Pitman, W. E.) 221–228 (American Geophysical Union, Washington DC, 1977).
30. Lanphere, M. S. & Reed, B. L. *Bull. geol. Soc. Am.* **84,** 3773–3782 (1972).
31. Vlaar, N. J. & Wortel, M. J. R. *Tectonophysics* **32,** 331–351 (1976).
32. Molnar, P. & Atwater, T. *Earth planet Sci. Lett.* **41,** 330–340 (1978).
33. Bluck, B. J., Halliday, A. N., Aftalion, M. & Macintyre, R. M. *Geology* **8,** 492–495 (1980).
34. Leggett, J., McKerrow, W. S. & Eales, M. H. *J. geol. Soc. Lond.* **136,** 755–770 (1979).
35. Johnson, M. R. W. & Frost, R. T. C. *Geol. Mijnb* **56,** 287–294 (1977).
36. Pankhurst, R. J. & Sutherland, D. S. in *Igneous Rocks of the British Isles* (ed. Sutherland, D.) (Wiley, New York, 1981).
37. Halliday, A. N., Aftalion, M. van Breemen & Jocelyn, J. in *The Caledonides of the British Isles*, 653–661 (Scottish Academic, Edinburgh, 1979).
38. Halliday, A. N., Stephens, W. S. & Harmon, R. S. *J. geol. Soc. Lond.* **137,** 329–348 (1980).
39. Pankhurst, R. J. in *Origin of Granite Batholiths, Geochemical Evidence* (eds Atherton, M. P. & Tarney, J.) 18–33 (Shiva, Copenhagen, 1979).
40. Raymond, L. A. & Swanson, S. E. *Nature* **385,** 317–319 (1980).
41. Raymond, L. A. *Earth planet. Sci. Lett.* **49,** 29–33 (1980).
42. Gilluly, J. *Geol. Soc. Am. Spec. Pap.* **62,** 7–18 (1955).
43. Armstrong, R. L. *Phil. Trans. R. Soc.* (in the press).
44. Harmon, R. S. & Halliday, A. N. *Nature* **283,** 21–25 (1980).
45. Blaxland, A. B., Aftalion, M. & van Breemen, O. *Scott. J. Geol.* **15,** 139–151 (1971).
46. Hamilton, P. J., O'Nions, R. K. & Pankhurst, R. J. *Nature* **287,** 279–283 (1980).
47. Doe, B. R. & Zartman, R. E. in *Geochemistry of Hydrothermal Ore Deposits* (ed. Barnes, H. L.) 22–27 (Wiley, New York, 1979).
48. Chapman, H. J. & Moorbath, S. *Nature* **268,** 41–42 (1977).
49. Brown, G. C. *Nature* **265,** 21–24 (1977).

Expression in *E. coli* of maize and wheat chloroplast genes for large subunit of ribulose bisphosphate carboxylase

Anthony A. Gatenby, Judith A. Castleton & Michael W. Saul
Plant Breeding Institute, Trumpington, Cambridge CB2 2LQ, UK

Synthesis of the large subunit polypeptide of ribulose bisphosphate carboxylase can be detected in Escherichia coli *cells containing the chloroplast genes from maize and wheat. The chloroplast DNA sequences contain a 'promoter' that, in* Escherichia coli, *initiates transcription and translation of the genes. This synthesis of large subunits has been substantially increased using the powerful λ promoter, P_L, to drive transcription of the maize gene.*

0028–0836/81/200117—05$01.00

ALTHOUGH much effort has recently been devoted to constructing plasmids in which the expression of mammalian genes can be directed from prokaryotic promoters[1–7], there have been no reports of the expression of plant genes in *Escherichia coli*. Chloroplast genes from higher plants would seem to be particularly favourable candidates for expression in bacteria because of the similarities already established between chloroplasts and prokaryotic organisms. Thus, in both there are 70S ribosomes containing 23S, 16S and 5S rRNA[8], there is strong sequence homology between maize chloroplast and *E. coli* 16S rRNA genes[9] and *N*-formylmethionyl-tRNA is used for the initiation of polypeptide chain synthesis[10]. Furthermore, chloroplast genes have been transcribed and translated successfully *in vitro* using *E. coli* cell-free extracts[11]. Accordingly, we have investigated chloroplast gene expression in bacteria using the maize and wheat genes for the large subunit polypeptide of ribulose bisphosphate carboxylase (RuBPCase).

Our choice of this gene was favoured by the extensive knowledge of both the structure and function of the gene and the enzyme. RuBPCase [3-phospho-D-glycerate carboxy-lyase (dimerizing), EC 4.1.1.39] holoenzyme is a major soluble protein in green plants and fixes CO_2 in the photosynthetic pentose phosphate pathway[12]. It is present in the stroma of chloroplasts as an oligomer with 16 subunits and a molecular weight (MW) of ~560,000 (ref. 13). Eight of these subunits are small (MW 12–18,000) and the remaining eight are large (MW 50–55,000). The large subunit contains the catalytic sites for the carboxylase and oxygenase reactions[12] but the function of the small subunit remains undetermined. The small subunits are nucleus encoded[14] and synthesized on cytoplasmic ribosomes as a precursor which passes through the chloroplast membrane and is processed to its final size[15,16]. It is assembled with large subunits that are encoded by chloroplast DNA[17] and synthesized on chloroplast ribosomes[18]. The DNA sequence of the maize large subunit gene[19] and a pea small subunit cDNA clone[20] have been established.

We describe here the synthesis of the large subunit polypeptide of RuBPCase in *E. coli*, apparently under the control of chloroplast sequences that function as promoters in this bacterium. The synthesis of the plant protein can also be manipulated using promoters from bacteriophage λ during lytic infection. The demonstration that chloroplast genes from higher plants are expressed in *E. coli* provides many new opportunities in plant and bacterial molecular biology.

Construction of λ and plasmid molecules containing RuBPCase large subunit gene sequences

The chloroplast chromosome of all characterized species contains a single copy of a gene for the large subunit of RuBPCase (RuBPCase LS), and the position of this gene has been mapped for maize[21] and wheat (C. M. Bowman, B. Koller, H. Delius and T. A. Dyer, personal communication). The gene from both species can be isolated intact on a fragment generated by digestion with the restriction endonuclease *Bam*HI; 4.3 kilobase pairs for maize and 9.7 for wheat. Both these fragments have been cloned directly into the *Bam*HI site of pBR322. The maize fragment has the additional advantage of being a suitable size for cloning into the *Bam*HI λ insertion vector NM570-BV2 (ref. 22). Our approach in the construction of λ and plasmid recombinants containing the RuBPCase LS sequences is detailed in Fig. 1 legend.

The orientation of the maize gene in λ was determined using the asymmetric location of *Eco*RI sites within the chloroplast DNA fragment[21] and the phage vector[22]. Clones were identified in which the RuBPCase LS gene is transcribed either from the *l*-strand of the phage chromosome using the promoter P_L(λ*Zm*B1^l), or from the *r*-strand using the promoter P'_R(λ*Zm*B1^r). The maize gene was transferred to the *Bam*HI site of pBR322 and the two possible orientations (p*Zm*B1A and p*Zm*B1B) distinguished by the asymmetric location of *Pst*I sites within the chloroplast DNA fragment[21] and the plasmid[23]. The plasmid p*Tac*39 contains the wheat RuBPCase large subunit gene on a fragment inserted into the *Bam*HI site of pBR322 and we derived the plasmid p*Tac*518 from p*Tac*39 by reversing the orientation of the chloroplast fragment (Fig. 1).

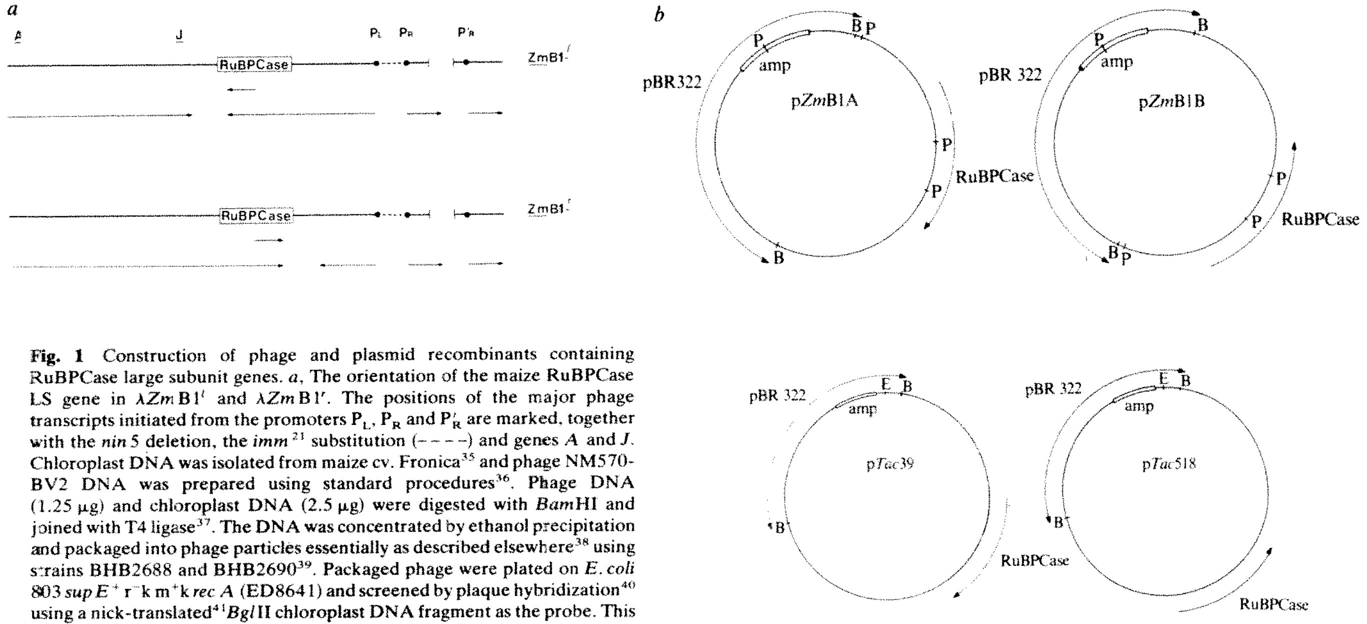


Fig. 1 Construction of phage and plasmid recombinants containing RuBPCase large subunit genes. *a*, The orientation of the maize RuBPCase LS gene in λ*Zm*B1^l and λ*Zm*B1^r. The positions of the major phage transcripts initiated from the promoters P_L, P_R and P'_R are marked, together with the *nin* 5 deletion, the *imm*[21] substitution (– – – –) and genes *A* and *J*. Chloroplast DNA was isolated from maize cv. Fronica[35] and phage NM570-BV2 DNA was prepared using standard procedures[36]. Phage DNA (1.25 μg) and chloroplast DNA (2.5 μg) were digested with *Bam*HI and joined with T4 ligase[37]. The DNA was concentrated by ethanol precipitation and packaged into phage particles essentially as described elsewhere[38] using strains BHB2688 and BHB2690[39]. Packaged phage were plated on *E. coli* 803 $supE^+\ r_k^-m_k^+\ recA$ (ED8641) and screened by plaque hybridization[40] using a nick-translated[41] *Bgl*II chloroplast DNA fragment as the probe. This fragment contains the complete RuBPCase LS gene sequence[21]. Positive clones were plaque purified and phage isolated from liquid lysates[36]. The orientation of the RuBPCase LS gene in the phage chromosome was deduced from the asymmetric distribution of *Eco*RI sites[21,22] and confirmed by transfer to nitrocellulose filters[42] and hybridization with the nick-translated *Bgl*II fragment. *b*, The orientation of the maize RuBPCase LS gene in plasmids p*Zm*B1A and p*Zm*B1B. The position of *Bam*HI (B) and *Pst*I (P) sites and the direction of transcription of the maize gene are indicated. λ*Zm*B1 DNA (2 μg) and pBR322 DNA (0.2 μg) were digested with *Bam*HI and joined with T4 ligase. The DNA was transformed into strain ED8641 and clones resistant to ampicillin and sensitive to tetracycline selected. Purified plasmids[43] with the RuBPCase LS gene in either of the two orientations were identified by the asymmetric distribution of *Pst*I sites[21,23] and confirmed by hybridization with the *Bgl*II fragment as described in *a*. *c*, The orientation of the wheat RuBPCase LS gene in plasmids p*Tac*39 and p*Tac*518. The position of *Bam*HI (B) sites and the direction of transcription of the wheat gene are indicated. The map of p*Tac*39 was produced from information provided by C. M. Bowman and T. A. Dyer. The wheat fragment was excised with *Bam*HI and re-joined with T4 ligase to obtain p*Tac*518 with the fragment in the opposite orientation.

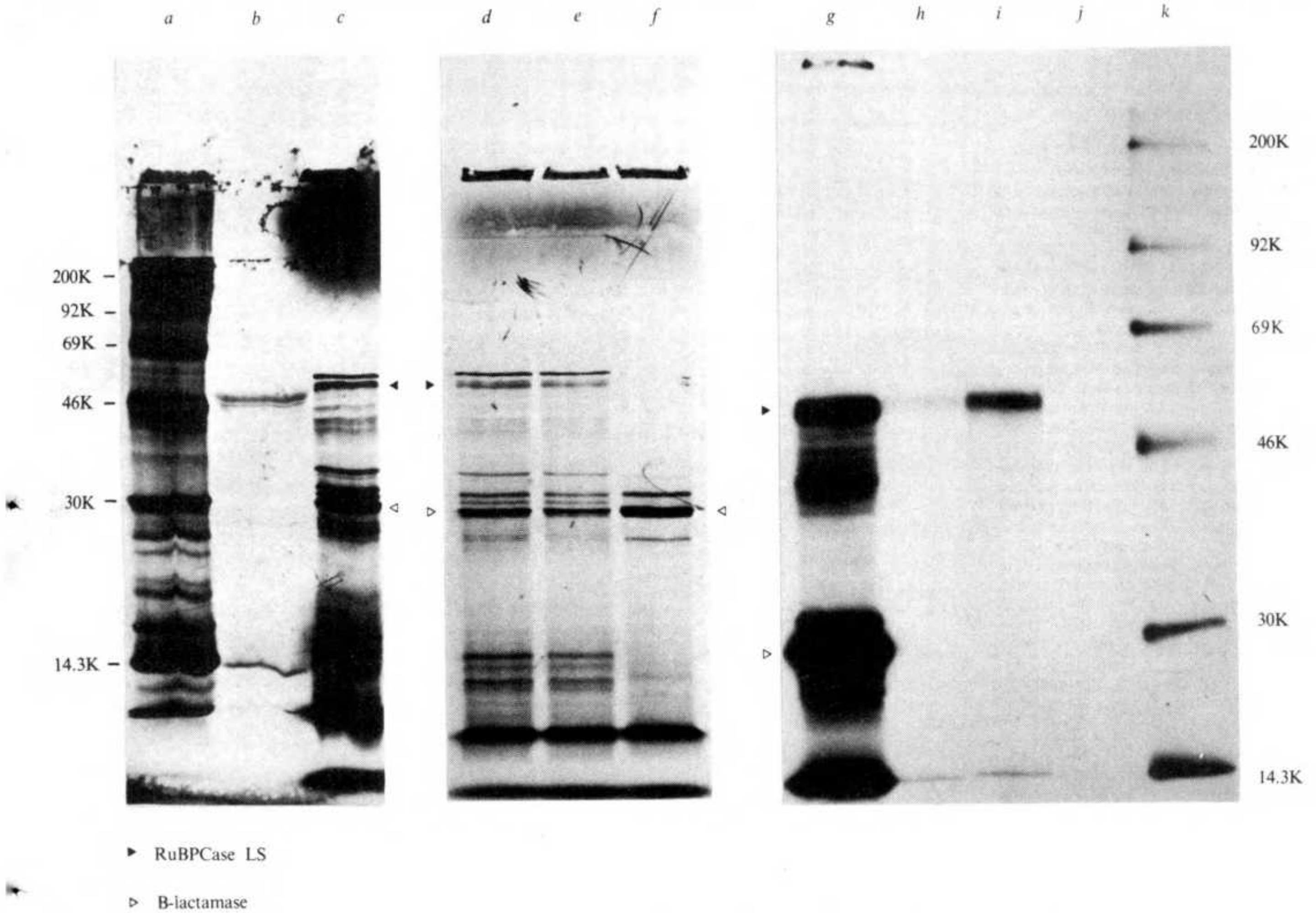


Fig. 2 Expression of RuBPCase LS in *E. coli* minicells. The plasmids pBR322, p*Tac* 39, p*Tac* 518, p*Zm* B1A and p*Zm* B1B were transformed into the minicell strain DS410[24]. Minicells were prepared from 200 ml of an overnight culture grown in brain heart infusion[24]. The minicells were preincubated for 30 min at 37 °C in minimal medium and then incubated for a further 60 min in the presence of 10–20 μCi L-[^{35}S] methionine (Amersham). Prewarmed L broth was added and after a further incubation for 15 min the minicells were collected by centrifugation and lysed by boiling in 100 μl of 2% (w/v) SDS for 2 min. The number of proteins synthesized was examined by SDS-polyacrylamide gel electrophoresis[44], followed by autoradiography of the dried gels. For immunoprecipitation the SDS minicell lysate was diluted with 1 ml of Triton buffer (50 mM Tris-HCl *p*H 7.8, 0.15 M NaCl, 2 mM EDTA, 1% (v/v) Triton X-100) and clarified at 20,000*g* for 10 min. To the supernatant was added 10 μl of either anti-RuBPCase serum or pre-immune serum. After incubation at 20 °C for 2 h the antigen–antibody complex was absorbed to protein A–Sepharose (equivalent to 5 mg dried Sepharose) by shaking gently for 1–2 h. The Sepharose was washed once with Triton buffer containing 1 M NaCl and three times with Triton buffer. The Sepharose pellet was resuspended in 20μl of SDS loading buffer and applied directly to a 10% polyacrylamide gel and electrophoresed in the presence of SDS[44]. The dried gels were examined by autoradiography or fluorography[29] at −80 °C. Wheat and maize RuBPCase were purified for immunization, and as standards, by chromatography on Sephadex G-75, Ultrogel AcA22 and DEAE-Sephadex A50[45] followed by sucrose gradient centrifugation[46]. Gels were calibrated with a ^{14}C-methylated protein mixture (Amersham, CFA.626). The molecular weight markers were lysozyme (14,300), carbonic anhydrase (30,000), ovalbumin (46,000), bovine serum albumin (69,000), phosphorylase *b* (92,000) and myosin (200,000). *a*, Molecular weight markers (long exposure); *b*, p*Tac* 39 incubated with anti-RuBPCase serum and protein A–Sepharose; *c*, total p*Tac* 39; *d*, total p*Tac* 39; *e*, total p*Tac* 518; *f*, total pBR322; *g*, total p*Zm* BlA; *h*, p*Zm* BlA incubated with a RuBPCase immunoabsorbent and eluted with 8 M urea[45]; *i*, p*Zm* BlA incubated with anti-RuBPCase serum and protein A–Sepharose; *j*, p*Zm* BlA incubated with pre-immune serum and protein A–Sepharose; *k*, molecular weight markers (short exposure).

In vivo expression of RuBPCase large subunit gene on plasmids

Plasmids with both orientations of the RuBPCase LS gene from maize (p*Zm* BlA and p*Zm* B1B) and wheat (p*Tac* 39 and p*Tac* 518) were transformed into the minicell-producing strain DS410[24]. Minicells containing each of the four plasmids and a pBR322 control were prepared and incubated in the presence of L-[^{35}S] methionine to label specifically proteins synthesized from the plasmids[24]. Autoradiography of extracts separated on SDS-polyacrylamide gels demonstrated that several additional polypeptides were synthesized when the recombinant plasmids were present in the minicells, compared with the pBR322 control (Fig. 2). One of these additional polypeptides had a MW of 52–53,000 and a similar mobility to the purified RuBPCase large subunit isolated from leaves (data not shown). The identity of this polypeptide was confirmed by incubating the *in vivo* labelled minicell extracts with anti-RuBPCase serum and recovering the antigen–antibody complex by absorption to protein A–Sepharose. Radioactive polypeptides eluted from the protein A–Sepharose were resolved by SDS-polyacrylamide gel electrophoresis mainly as a major band of MW 52–53,000, although traces of other polypeptides could be detected after prolonged fluorography (Fig. 2). The RuBPCase LS polypeptide was not recovered if pre-immune serum was incubated with the minicell extracts. The slight difference in mobility between the immunoabsorbed wheat polypeptide and the corresponding polypeptide in the total minicell extract (Fig. 2, tracks *b*, *c*) is probably due to distortion in the gel by an unlabelled polypeptide present in large amounts in this part of the gel, which we suspect to be the heavy chain of IgG that is absorbed to the protein A–Sepharose.

The number and relative amounts of polypeptides synthesized from the RuBPCase LS plasmids were independent of the orientation of the chloroplast fragments in pBR322. Tran-

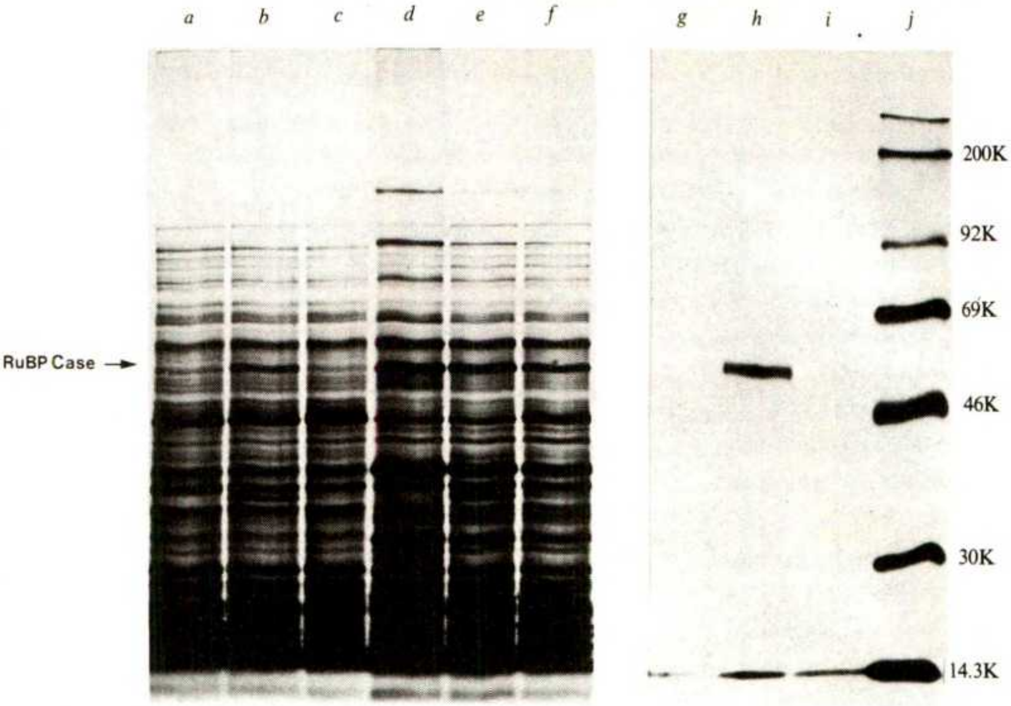


Fig. 3 Expression of maize RuBPCase LS during lytic infection with phage recombinants. *E. coli S*159 was grown in minimal medium with maltose and UV light-irradiated essentially as described by Jaskunas *et al.*[25]. 50 µl (5×10^7 bacteria) were mixed on ice with 5 µl (1×10^8 plaque-forming units) of phage suspension. For cells labelled early during infection the samples were incubated at 37 °C for 3 min and then diluted with 200 µl of prewarmed M9 salts, 1 mM $MgSO_4$, 0.3% (w/v) maltose, 0.04 µg methionine per ml, containing 10 µCi of L-[^{35}S]methionine. After 10 min the cells were chased with 30 µl of 2 mg ml^{-1} methionine for 1 min. For cells labelled late during infection the samples were incubated at 37 °C for 12 min, diluted with the isotope solutions as above and incubated for 10 min, and finally chased as before. Protein synthesis was terminated by the addition of 1 ml of cold acetone at −20 °C for 30 min. Precipitated proteins were collected by centrifugation and dried. Total protein synthesis was examined by dissolving the pellet in SDS loading buffer, resolving the polypeptides by electrophoresis on 10% polyacrylamide gels[44] and autoradiography of the dried gels. For immunoabsorption the precipitated protein was boiled for 2 min in 100 µl of 2% (w/v) SDS, cooled to room temperature and diluted with 1 ml of Triton buffer (Fig. 2). Conditions for incubating with antiserum and protein A–Sepharose were identical to those for Fig. 2. The immunoabsorbed protein was detected by fluorography[28] at −80 °C. Early labelling with: *a*, λNM 570–BV2; *b*, $\lambda Zm\text{B}1^{l}$; *c*, $\lambda Zm\text{B}1^{r}$. Late labelling with: *d*, λNM 570–BV2; *e*, $\lambda Zm\text{B}1^{l}$; *f*, $\lambda Zm\text{B}1^{r}$. Immunoabsorbed polypeptides after early infection with: *g*, λNM 570–BV2; *h*, $\lambda Zm\text{B}1^{l}$; *i*, $\lambda Zm\text{B}1^{r}$. The MW markers (*j*) were identical to those of Fig. 2.

scription and translation from sites within the chloroplast fragments seems the most likely explanation of this expression, in view of the absence of an orientation effect on the synthesis of maize and wheat RuBPCase LS. What is known of pBR322 transcription suggests that the level of expression of a gene using external promoters would be very different if inserted in opposite orientations in the *Bam*HI site.

To estimate the proportion of RuBPCase LS synthesized compared with β-lactamase, which is encoded on the same plasmid, we used scanning densitometry of the autoradiographs. After adjustment for the methionine content of RuBPCase LS and β-lactamase, the amount of RuBPCase LS synthesized is equivalent to approximately a 1 : 3 molar ratio.

Lytic expression of maize RuBPCase LS

To increase the synthesis of the large subunit of maize RuBPCase in *E. coli* we used P_L-initiated transcription to drive expression of the chloroplast gene. Phage-directed protein synthesis was examined by infecting UV light-irradiated *S*159 cells in the presence of L-[^{35}S]methionine[25] and resolving the products by polyacrylamide gel electrophoresis (Fig. 3). Incubation of lysates with anti-RuBPCase serum and immunoabsorption to protein A–Sepharose were used to verify the identity of the large subunit (Fig. 3).

Lytic expression of maize RuBPCase LS is strongly dependent on the orientation of the gene in the phage chromosome and on transcription from λ promoters. The kinetics of synthesis of the polypeptide (Fig. 4) demonstrate that when the chloroplast gene is transcribed from the *l*-strand of λ, RuBPCase LS accumulates rapidly in the first 10 min post-infection and accounts for 0.5–1.0% of total protein synthesis at this stage; RuBPCase LS synthesis then decreases. This pattern of expression is entirely consistent with the synthesis of RuBPCase LS as an early protein by P_L-initiated transcription, followed by repression of P_L by the phage *cro* gene product[26]. When the

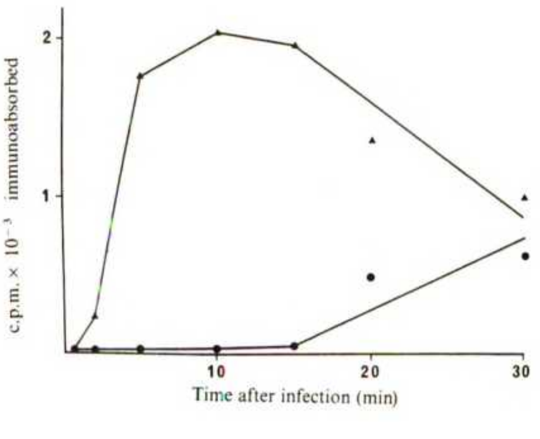


Fig. 4 Kinetics of expression of maize RuBPCase LS gene in *E. coli* cells infected with $\lambda \text{ZmB}1^{l}$(▲) or $\lambda Zm\text{B}1^{r}$(●). UV light-irradiated *S*159 cells were prepared[25]. 300 µl (3×10^8 bacteria) were mixed on ice with 30 µl (6×10^8 plaque-forming units) of phage suspension. After incubation at 37 °C for 3 min, 1.2 ml of prewarmed M9 salts, 1 mM $MgSO_4$, 0.3% (w/v) maltose, 0.05 µg L-methionine and 60 µCi L-[^{35}S]methionine were added. Duplicate 100-µl samples were withdrawn at the times indicated and added to 1 ml of cold acetone. After 30 min at −20 °C the proteins were collected by centrifugation, dried and boiled in 100 µl of 2% (w/v) SDS for 2 min. Triton buffer (1 ml) (Fig. 2) was added and after centrifuging at 20,000*g* for 10 min, an excess of anti-RuBPCase serum (20 µl) was added to the supernatant and incubated at 4 °C overnight. The antigen–antibody complex was absorbed to excess protein A–sepharaose (equivalent to 10 mg dried Sepharose) by shaking gently for 2 h at 20 °C. The Sepharose was washed three times with 1 – ml volumes of 1 M NaCl and once with 0.15 M NaCl, both in Triton buffer. The absorbed protein was eluted with four 250-µl aliquots of 50 mM Tris-HCl *p*H 8.0, 8 M urea, 0.2 M NaCl and 100 µg ml^{-1} bovine serum albumin. The four aliquots were combined and protein precipitated with an equal volume of cold 10% (w/v) trichloroacetic acid. The precipitates were collected on GF/C filters, washed three times with 5% (w/v) trichloroacetic acid, once with ethanol and dried. The filters were processed by liquid scintillation counting. The results were corrected for background by using *S*159 cells infected with the vector NM570-BV2 in the presence of L-[^{35}S]methionine and processed as above.

orientation of the chloroplast gene is reversed so that phage-initiated transcription is on the *r*-strand, the pattern of synthesis follows that expected of a late phage protein[27] and occurs after 15–20 min post-infection. The maize chloroplast fragment was cloned into the sites $sbam\lambda$-2 and $sbam\lambda$-3 (ref. 22), thus eliminating the phage *b* region and its transcription termination sites. This explains the observed transcription of the chloroplast gene from phage promoters regardless of the orientation of the gene.

To confirm the function of the phage promoters on the enhanced expression of the maize RuBPCase LS gene we infected a homoimmune λ lysogen, $S159(imm^{21})$, with $\lambda Zm\mathrm{B}1^{l}$ and $\lambda Zm\mathrm{B}1^{r}$. The lysogen synthesizes phage repressor to block the major leftward and rightward transcriptions[28] during infection with the λ recombinants, thus allowing the effect of phage transcription on synthesis of RuBPCase LS to be determined. The fact that we were unable to detect any synthesis of RuBPCase LS by fluorography[29] of immunoabsorbed extracts from the infected lysogens indicates that the enhanced expression of the chloroplast gene is dependent on λ transcription. Expression of the RuBPCase gene in infected lysogens using the endogenous chloroplast promoter was also not detectable, probably due to the low copy (two) of the template introduced during infection.

Discussion

We have demonstrated that the maize and wheat chloroplast genes for the large subunit of RuBPCase are expressed in *E. coli* to give a product of the correct size. This expression is apparently directed on plasmids by transcription initiated from chloroplast sequences that function as promoters. The maize RuBPCase LS gene and flanking regions have been sequenced and were reported to lack a 'classical' prokaryotic promoter sequence[19]. However, scrutiny of the published sequence[19] does reveal structural features similar to those observed in prokaryotic promoters. The sequences of special interest lie between −80 and −125 nucleotides before the start of the codon for the *N*-terminal methionine residue. The region includes the sequence TTGATA (base pairs −119 to −114) which has a high degree of homology with the conserved TTGACA sequence frequently found (around −35 base pairs) before the initiation point of mRNA transcription in prokaryotes[30]. Such a sequence is implicated in the initial recognition of the promoter site by *E. coli* RNA polymerase[30]. This sequence in maize is followed by the sequence TAGATT (base pairs −92 to −87) in a similar position and has significant homology to a Pribnow box[31]. Additional regions of less conserved homology are present on either side of the putative Pribnow box and the TTGATA sequence in maize when compared with a prototype prokaryotic promoter sequence[30]. A typical prokaryotic ribosome binding site GGAGG[9] is centred at −8 base pairs before the N-terminal methionine codon. Koller and Dyer (personal communication) have demonstrated by electron microscopy that *E. coli* RNA polymerase holoenzyme binds specifically to a region on the maize fragment similar to that possessing the sequences described above.

The maize chloroplast DNA sequence therefore possesses features that would lead one to suspect that expression of this plant gene in *E. coli* was possible and our results confirm this expectation. The amount of RuBPCase LS synthesized compared with β-lactamase indicates that expression of the chloroplast gene from a chloroplast promoter in *E. coli* is not insignificant. The ability of chloroplast sequences to initiate transcription has been investigated further by ligating *Sau*3A-generated chloroplast fragments from the wheat RuBPCase LS plasmid p*Tac*39 into the *Bam*HI site of the promoter-deleted β-galactosidase plasmid pMC489[32] and screening recombinants for the synthesis of the β-galactosidase fragment. Clones have been isolated that synthesize the β-galactodase fragment when chloroplast sequences are placed at the 5′ end of the *E. coli* gene, suggesting that initiation of transcription of β-galactosidase mRNA occurs from a chloroplast promoter sequence[33].

The expression of chloroplast genes in *E. coli*, if a general phenomenon, will be advantageous for several reasons. (1) The synthesis *in vivo* of chloroplast-encoded antigens in bacterial colonies or phage plaques will allow solid-phase radioimmunoassay[34] to be used for the rapid identification of clones of interest, avoiding the time-consuming method of *in vitro* coupled transcription–translation systems to identify clones[11]. (2) It should also make possible the identification of some chloroplast genes by complementation of *E. coli* mutants. (3) *In vitro* and *in vivo* mutagenesis and modification of cloned chloroplast genes can be accomplished and the results of such changes on the expression of the gene and the nature of the product can be monitored in bacterial cells. (4) The use of *E. coli* to synthesize chloroplast proteins will be of value for the purification of proteins present only in small amounts in the plant cell, and in the purification of the subunit components of oligomeric enzymes.

Our demonstration that the level of synthesis of RuBPCase large subunits can be increased by using phage promoters should facilitate examination of the assembly mechanism of the RuBPCase holoenzyme and the catalytic function of isolated large subunits that have not been denatured during preparation.

We thank R. B. Flavell for continual encouragment and interest, C. M. Bowman and T. A. Dyer for p*Tac*39, the maize *Bgl*II fragment and unpublished results, J. C. Gray for raising antiserum and K. and N. Murray for instruction in the use of λ and for providing most of the phage and bacterial strains used. Financial support was provided by the National Research Development Corporation.

Received 1 December 1980; accepted 5 March 1981.

1. Derynck, R. *et al. Nature* **287,** 193–197 (1980).
2. Nagata, S. *et al. Nature* **284,** 316–320 (1980).
3. Goeddel, D. V. *et al. Nature* **281,** 544–548 (1979).
4. Seeburg, P. H. *et al. Nature* **276,** 795–798 (1978).
5. Shine, J. *et al. Nature* **285,** 456–461 (1980).
6. Erwin, C. R., Maurer, R. A. & Donelson, J. E. *Nucleic Acids Res.* **8,** 2537–2546 (1980).
7. Chang, A. C. Y. *et al. Nature* **275,** 617–624 (1978).
8. Bedbrook, J. R. & Kolodner, R. *A.·Rev. Pl. Physiol.* **30,** 593–620 (1979).
9. Schwarz, Zs. & Kossel, H. *Nature* **283,** 739–742 (1980).
10. Schwartz, J. H., Meyer, R., Eisenstadt, J. M. & Brawerman, G. *J. molec. Biol.* **25,** 571–574 (1967).
11. Whitfield. P. R. & Bottomley, W. *Biochem. Int.* **1,** 172–178 (1980).
12. Jensen, R. G. & Bahr, J. T. *A. Rev. Pl. Physiol.* **28,** 379–400 (1977).
13. Baker, T. S., Eisenberg, D. & Eiserling, F. *Science* **196,** 293–295 (1977).
14. Kawashima, N. & Wildman, S. G. *Biochim. biophys. Acta* **262,** 42–49 (1972).
15. Highfield, P. E. & Ellis, R. J. *Nature* **271,** 420–424 (1978).
16. Chua, N. H. & Schmidt, G. W. *Proc. natn. Acad. Sci. U.S.A.* **75,** 6110–6114 (1978).
17. Coen, D. M., Bedbrook, J. R., Bogorad, L. & Rich, A. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5487–5491 (1977).
18. Blair, G. E. & Ellis, R. J. *Biochim. biophys. Acta* **319,** 223–234 (1973).
19. McIntosh, L., Poulsen, C. & Bogorad, L. *Nature* **288,** 556–590 (1980).
20. Bedbrook, J. R., Smith, S. M. & Ellis, R. J. *Nature* **287,** 692–697 (1980).
21. Bedbrook, J. R. *et. al. J. biol. Chem.* **254,** 905–910 (1979).
22. Klein, B. & Murray, K. *J. molec. Biol.* **133,** 289–294 (1979).
23. Sutcliffe, J. G. *Nucleic Acids Res.* **5,** 2721–2728 (1978).
24. Dougan, G. *et al. J. Bact.* **138,** 48–54 (1979).
25. Jaskunas, S. R., Lindahl, L., Nomura, M. & Burgess, R. R. *Nature* **257,** 458–462 (1975).
26. Johnson, A., Meyer, B. J. & Ptashne, M. *Proc. natn. Acad. Sci. U.S.A.* **75,** 1783–1787 (1978).
27. Hendrix, R. W. in *The Bacteriophage Lambda* (ed. Hershey, A. D.) 355–370 (Cold Spring Harbor Laboratory, New York, 1971).
28. Ptashne, M. in *The Operon* (eds Miller, J. H. & Reznikoff, W. S.) 325–343 (Cold Spring Harbor Laboratory, New York, 1978).
29. Chamberlain, J. P. *Analyt. Biochem.* **98,** 132–135 (1979).
30. Rosenberg, M. & Court, D. *A. Rev. Genet.* **13,** 319–353 (1979).
31. Pribnow, D. *J. molec. Biol.* **99,** 419–443 (1975).
32. Casadaban, M. J. & Cohen, S. N. *J. molec. Biol.* **138,** 179–207 (1980).
33. Saul, M. W. (in preparation).
34. Broome, S. &. Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **75,** 2746–2749 (1978).
35. Kolodner, R. & Tewari, K. K. *Biochim. biophys. Acta.* **402,** 372–390 (1975).
36. Murray, N. E., Brammer, W. J. & Murray, K. *Molec. gen. Genet.* **150,** 53–61 (1977).
37. Murray, N. E., Bruce, S. A. & Murray, K. *J. molec. Biol.* **132,** 493–505 (1979).
38. Enquist, L. & Sternberg, N. *Meth. Enzym.* **68,** 281–298 (1979).
39. Hohn, B. *Meth. Enzym.* **68,** 299–309 (1979).
40. Benton, W. D. & Davis, R. W. *Science* **196,** 180–182 (1977).
41. Maniatis, T., Jeffrey, A. & Kleid, D. G. *Proc. natn. Acad. Sci. U.S.A.* **72,** 1184–1188 (1975).
42. Southern E. M. *J. mol. Biol.* **98,** 503–517 (1975).
43. Clewell, D. B. & Helsinki, D. R. *Proc. natn. Acad. Sci. U.S.A.* **62,** 1159–1166 (1969).
44. Laemmli, U. K. *Nature* **227,** 680–685 (1970).
45. Gatenby, A. A. & Cocking, E. C. *Pl. Sci. Lett.* **13,** 171–176 (1978).
46. Goldthwaite, J. J. & Bogorad, L. *Analyt. Biochem.* **41,** 57–66 (1971).

Genetic analysis of K88-mediated adhesion of enterotoxigenic *Escherichia coli*

Michael Kehoe[*], Richard Sellwood[†], Patricia Shipley[‡] & Gordon Dougan[*]

[*] Department of Microbiology, Moyne Institute, Trinity College, Dublin 2, Ireland
[†] Institute for Research on Animal Diseases, Compton, Newbury, Berkshire, UK
[‡] Department of Microbiology, Virginia Commonwealth University, Richmond, Virginia 23298, USA

Four cistrons (adh) *involved in the expression of the K88 adhesion system have been identified and mapped. Three of these* (adh *A*, adh *B and* adh *C*) *are located in a single operon* (*I*) *whereas the fourth* (adh *D*) *is expressed from a separate promoter* (*operon II*). *The polypeptides encoded by these cistrons have been identified and their role in the formation and regulation of K88 fimbriae* (*pili*) *is discussed.*

THE ability to adhere specifically to mammalian epithelial cells is an important virulence characteristic of many pathogenic bacteria[1,2]. Adhesion to host tissues is often the first step in the establishment of an infection[3], thus elucidation of the mechanisms promoting adhesion could lead to the development of novel anti-adhesive drugs and improved vaccines. Many enterotoxigenic strains of *Escherichia coli* adhere specifically to the epithelial cells of the host's small intestine, where they secrete enterotoxins which are an important cause of diarrhoea in piglets, calves and human infants[4-7].

The adhesiveness of these *E. coli* strains is host specific and often associated with characteristic fimbrial proteinaceous antigens on the bacterial cell surface[8,9]: K88 and 987P in porcine strains[10,11], K99 in bovine strains[12], and the colonizing-factor antigens in human strains[13]. Each of these antigens is encoded by large (~70 kilobase pair) plasmids[7,13]. An 11.5-kilobase pair DNA fragment encoding the K88 serotype *ac* determinant has been cloned[14,15] into the multicopy vector plasmid pBR322 (ref. 16). The K88$^+$ hybrid plasmid was termed pPS002. To provide conclusive evidence for the role of the K88 antigen in adhesion and to determine the genetic basis of its expression, a number of deletion and transposon Tn*5* insertion mutants of pPS002 were isolated and analysed.

Deletion mutants of pPS002 generated *in vitro*

E. coli K12 strain C600 harbouring pPS002 readily agglutinates in anti-K88 sera (K88$^+$), causes mannose-resistant haemagglutination (MRHA$^+$) of guinea pig erythrocytes, carries K88 fimbriae on its surface and adheres strongly to isolated porcine brush border cells *in vitro* (Adh$^+$) (Table 1, Fig. 1). C600 has none of these properties. All the above tests are normally used to identify cells carrying wild-type K88*ac* plasmids.

The smallest deletion of pPS002 that was K88$^+$, MRHA$^+$, Adh$^+$ is pMK005 (Table 1, Fig. 2). This deletion mutant differs from pPS002 in that a 5.0-kilobase pair *Eco*RI fragment (E$_2$) is deleted. The remaining 6.5 kilobase pairs of cloned DNA sequence are sufficient to code for the K88$^+$, MRHA$^+$, Adh$^+$ phenotypes.

The deletion mutants pMK001, pMK002, and pMK006 are K88$^-$, MRHA$^-$ and Adh$^-$. Restriction endonuclease analysis of pMK001 DNA showed that this mutation resulted from a deletion extending from near the *Eco*RI site in pBR322 into the cloned DNA sequences and retains only ~1.6 kilobase pairs of the original cloned fragment (Fig. 2). Mutants pMK002 and pMK006 differ from pMK005 in that pMK002 has lost the 1.73-kilobase pair *Eco*RI fragment E$_3$ and pMK006 has lost an extensive part of it. Thus at least part of the sequences encoding the K88$^+$, MRHA$^+$, Adh$^+$ phenotypes are located within the 1.73-kilobase pair *Eco*RI fragment E$_3$.

To determine whether E$_3$ contains sufficient information to express the K88 antigen, the fragment was cloned separately into the single *Eco*RI site on the cloning vector pAC184 (ref. 17). The resulting hybrid plasmid, pMK007, was K88$^-$, MRHA$^-$ and Adh$^-$ (Table 1). These defects could not be

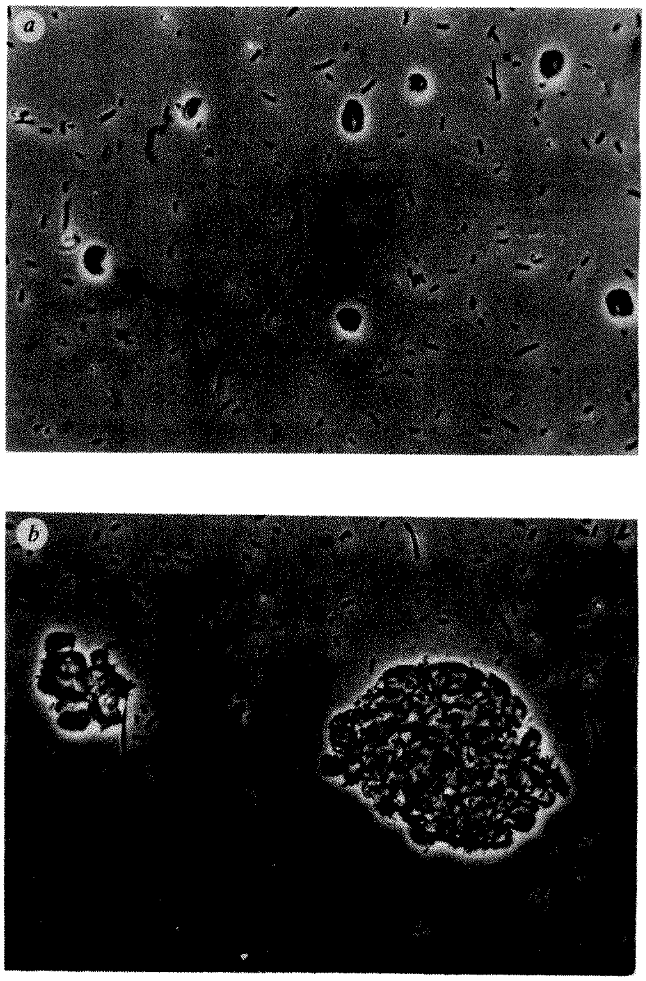


Fig. 1 Adhesion of K88$^+$ Adh$^+$ strains to porcine brush border cells. The adhesion tests were carried out in the presence of 0.5% (w/v) mannose, as described elsewhere[25]. Adhesion was scored from − to + + + +. *a*, Pig brush border cells (arrowed) plus the K88$^-$ Adh$^-$ host strain C600; *b*, pig brush border cells plus C600 harbouring the K88$^+$ Adh$^+$ plasmid, pPS002.

0028–0836/81/200122—05$01.00

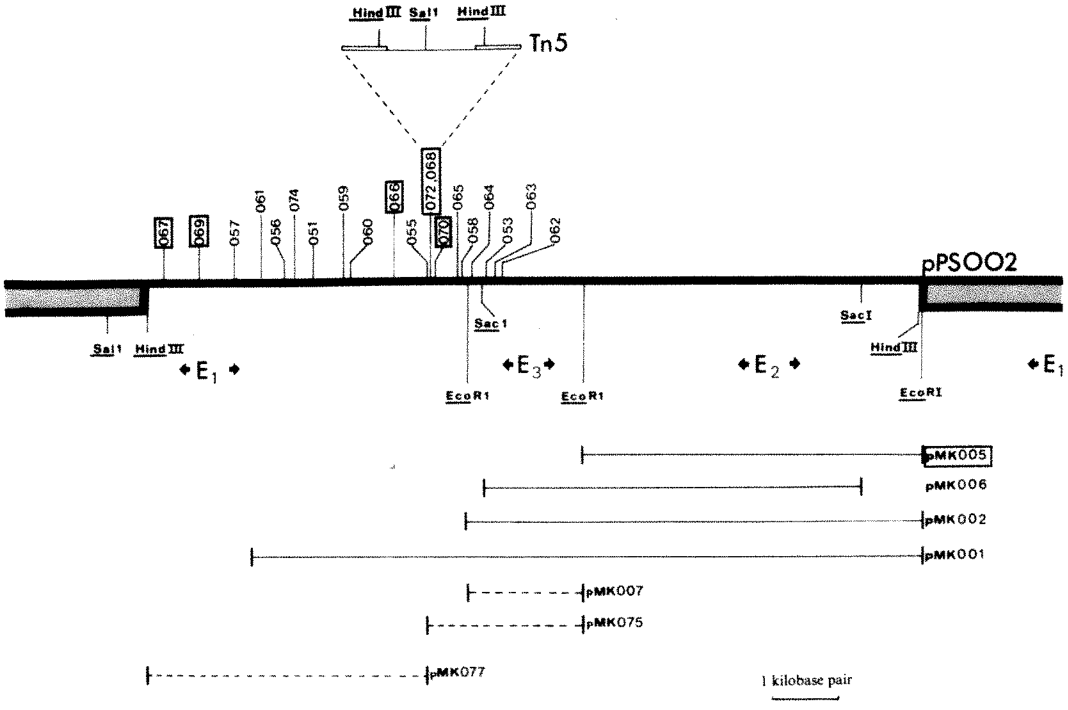


Fig. 2 Mutants of pPS002. The central solid line represents the cloned DNA sequences in pPS002; the wider shaded boxes at each end represent vector (pBR322) DNA sequences. The cleavage sites for several restriction endonucleases are indicated below the line. The three *Eco*RI fragments are indicated by the letters E_1, E_2 and E_3 in order of decreasing size. The lines beneath correspond to the DNA sequences which have been deleted in the mutants indicated; broken lines correspond to the DNA sequences which have been cloned separately into another vector to generate the hybrid plasmids indicated. The deletion mutant pMK001 was isolated after transforming C600 with linear pPS002 DNA generated by partial cleavage with *Eco*RI. The mutants pMK002, pMK005 and pMK006 were generated by deleting particular restriction enzyme fragments *in vitro*. The hybrid plasmid pMK007 was constructed by cloning *Eco*RI fragment E_3 into the vector pAC184. pMK075 was constructed by recircularizing the largest *Hin*dIII fragment of the pMK005::Tn*5* mutant, pMK070. Cloning of a different *Hin*dIII fragment from pMK070 into pAC184 generated the hybrid pMK077. In addition to the pPS002 DNA sequences indicated, the hybrid plasmids pMK075 and pMK077 contain 1.16 kilobase pairs of Tn*5* DNA sequences. The sites of the Tn*5* insertions are indicated by the perpendicular lines above pPS002. The structure of Tn*5* is shown above mutant pMK068[26]. The positions of the Tn*5* insertions were mapped by a series of *Sal*I and *Eco*RI, *Hin*dIII double digests. Mutants whose numbers are enclosed in boxes (for example pMK068) are K88$^+$Adh$^+$; the remaining mutants are K88$^-$Adh$^-$. Plasmid DNA was isolated as described elsewhere[27]. The conditions for restriction enzyme digests and the T_4 ligase reactions were as recommended by the manufacturers. Transformation of C600 was carried out as described elsewhere[28] and agarose gel electrophoresis as described in ref. 29. The Tn*5* insertions into pMK005 were isolated by infecting a Su$^-$ host harbouring pMK005 with the phage λ 467 b221 cI_{857} O_{29} P_{80} *rex*::Tn*5* at a multiplicity of ~0.5. This phage cannot lysogenize, nor can it replicate in a Su$^-$ host. Phage and host were mixed and incubated for 2–3 h at room temperature to allow transposition of Tn*5* and expression of the kanamycin resistance (Km^r) genes. Dilutions were plated onto nutrient agar containing kanamycin. The plasmid DNA from ~2,000 pooled Km^r colonies was used to transform C600, selecting for ampicillin resistance Km^r transformants. Only one pMK005::Tn*5* transformant from each infection experiment was used for further study.

complimented when pMK007 was introduced into C600 harbouring pMK002, indicating either that a gene or an operon involved in the expression of all three phenotypes spans the *Eco*RI site within the cloned DNA sequences in pMK005.

Tn*5* insertion mutants of pMK005

To map in more detail the DNA sequences encoding the K88$^+$, MRHA$^+$, Adh$^+$ phenotypes, a series of independent Tn*5* insertion mutants in plasmid pMK005 was isolated and mapped (Fig. 2). Twenty-three insertions within the cloned DNA sequences were chosen for further study (see Table 1); 19 of the pMK005::Tn*5* plasmids were phenotypically K88$^-$, MRHA$^-$ and Adh$^-$. These mutations spanned an extensive region (~5.5 kilobase pairs) of the cloned DNA sequences. Five of the Tn*5* insertions mapping within the cloned sequences were still K88$^+$, MRHA$^+$, Adh$^+$; two of these (pMK069 and pMK067) mapped close to one end of the sequence whereas three (pMK072, pMK070 and pMK068) were clustered in a small region in the centre. This cluster of K88$^+$, MRHA$^+$, Adh$^+$ Tn*5* insertions effectively divides pMK005 into two regions which accommodate K88$^-$, MRHA$^-$, Adh$^-$ Tn*5* insertion mutations (see Fig. 2). As insertion of Tn*5* into a proximal gene in an operon exerts a strong polar effect (~99%) on the expression of distal genes[18], this suggests that the genes involved in the expression of the K88 determinant are organized into at least two distinct operons.

Cells harbouring one of the insertion mutants, pMK066, were found to be K88$^+$ and Adh$^+$ but were MRHA$^-$. It has been assumed that the reactions involved in the adhesion of K88$^+$ enterotoxigenic *E. coli* strains to pig brush border cells and to guinea pig red blood cells are identical but the mutant pMK066 indicates that the situation may be a little more complex.

Polypeptides encoded by pMK005 and its deletion mutants in minicells

The two regions of pMK005 DNA which accommodate Tn*5* insertions altering the expression of the K88$^+$ Adh$^+$ MRHA$^+$ phenotypes are large enough to encode for polypeptides of molecular weight (MW) up to at least 138,000. The *E. coli* minicell system[19,20] was used to detect pMK005-encoded polypeptides (Table 1, Fig. 3). In addition to the vector (pBR322)-encoded β-lactamase polypeptides, the K88$^+$ Adh$^+$ MRHA$^+$ plasmid, pMK005, expressed at least four polypeptides of MWs 70,000, 29,000, 23,500 and 17,000 (Fig. 3*a*). Minicells harbouring the extensive deletion mutant, pMK001, failed to express any of these four polypeptides (Fig. 3*d*) indicating that they are encoded within the cloned DNA sequences of pMK005. The 23,500-MW polypeptide which has previously been shown to be the K88 antigen is expressed at a very high rate, forming up to 4% of the total cell protein[14].

The deletion mutants pMK002 and pMK006 expressed high levels of the 70,000, 29,000 and 17,000-MW polypeptides in minicells but failed to produce detectable quantities of the 23,500-MW K88 antigen (Fig. 3*b, c*). Thus, the cistrons encoding the former three polypeptides are located within the 4.8 kilobase pair cloned DNA sequences retained by pMK002 (Fig. 2) whereas at least part of the DNA sequences required for the expression of the K88 antigen are located within the

Table 1 Characteristics of pPS002, pMK005 and their mutant derivatives

Mutants	Agglutination with anti-K88 serum	MRHA	Adhesion to pig brush border cells *in vitro*	Polypeptides detected in minicells (MW) 70,000	29,000	23,500	17,000
C600 host strain	–	–	–	–	–	–	–
pPS002, pMK005, pMK067 pMK069, pMK070, pMK072, pMK068	+	+	++++	+	+	+	+
pMK001	–	–	–	–	–	–	–
pMK002, pMK006, pMK062, pMK063, pMK064, pMK065, pMK058, pMK053	–	–	–	+	+	–	+
pMK055	–	–	–	+	+	–	–
pMK066	+	–	++++	+	–	–(IG+)	–
pMK060, pMK059, pMK051, pMK074, pMK056, pMK061, pMK057	–	–	–	–	–	–(IG+)	–
pMK075	–	–	–	–	–	–	–
pMK077	–	–	–	+	+	–	+
pMK075 + pMK077	+	–	++++	+	+	+	+

Slide agglutination tests using anti-K88 serum were done. Serum was prepared by inoculating rabbits with K88 fimbrial subunit, purified as described by Mooi and de Graff[22]. The anti-K88 serum was dialysed and concentrated as described in ref. 23 and absorbed against C600 before use. Mannose-resistant haemagglutination (MRHA) of guinea pig erythrocytes was tested in the presence of 0.5% mannose as described in ref. 24. *In vitro* adhesion tests were carried out in the presence of 0.5% mannose as described elsewhere[25]. The degree of adhesion was scored from – to ++++. IG+ indicates that in these mutants low levels of the 23,500-MW K88 antigen could be detected by the immunoprecipitin gel overlay technique described in Fig. 4 legend.

1.7-kilobase pair *Eco*RI fragment, E_3. Minicells harbouring the pAC184–E_3 hybrid plasmid, pMK007, failed to express the K88 antigen (Fig. 3*e*).

Polypeptides expressed by pMK005::Tn*5* mutants in minicells

In addition to polypeptides expressed from pMK005 DNA sequences, all the pMK005::Tn*5* mutants expressed four polypeptides in minicells which are known to be encoded by the transposon[21]. The K88$^+$ Adh$^+$ MRHA$^+$ pMK005::Tn*5* mutants (pMK067, pMK068, pMK069, pMK070 and pMK072) expressed high levels of the 70,000, 29,000, 23,500 and 17,000-MW polypeptides described above in minicells (Fig. 3*s*). Six of the K88$^-$ Adh$^-$ MRHA$^-$ pMK005::Tn*5* mutants (pMK053, pMK058, pMK062, pMK064 and pMK065) expressed high levels of the 70,000, 29,000 and 17,000-MW polypeptides but failed to express detectable levels of the 23,500-MW K88 antigen (Fig. 3*g*–*i*, *k*). If the 23,500-MW polypeptide was being expressed in reduced quantities by these mutants it might not be distinguishable from the 24,000-MW Tn*5*-encoded neomycin phosphotransferase band on the fluorographs, nor would it be readily detected by Coomassie blue staining of SDS polyacrylamide gels of total cell protein. However, the immunoprecipitin gel overlay technique described in Fig. 4 legend also failed to detect any significant production of the K88 antigen by the six mutants (data not shown). The Tn*5* insertions in all six of these mutants map adjacent to each other within a ~650-base pair region which spans the *Eco*RI site separating *Eco*RI fragments E_1 and E_3 (Fig. 2). Thus the position of the cistron (*adh* D) encoding the 23,500-MW K88 antigen can be mapped (Fig. 5). The direction of transcription of this cistron is tentatively suggested by the observation that mutants pMK062 and

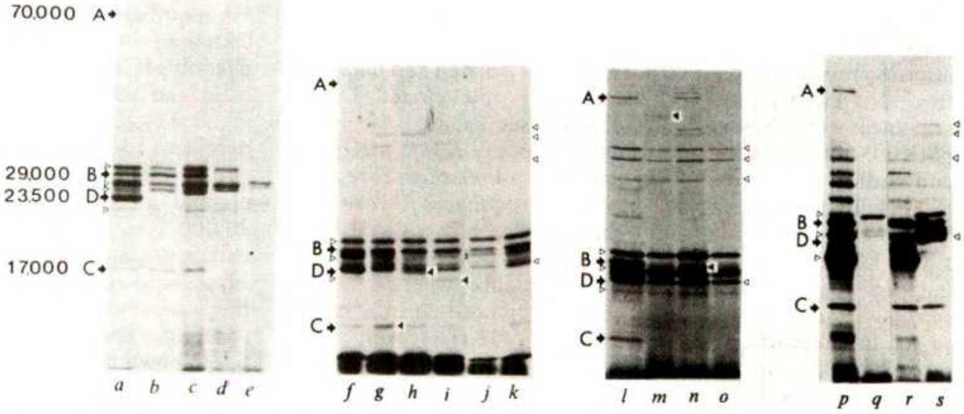


Fig. 3 Fluorographs of polypeptides expressed in minicells by pMK005 and its mutant derivatives. Where a number of different mutants produce the same pattern of polypeptides only representative mutants are shown. The pMK005-encoded polypeptides are indicated by solid black arrows at the side of the diagram; open arrows indicate vector-encoded polypeptides, that is, β-lactamase and its precursors in the case of pBR322-based hybrid plasmids (for example, *a*) or chloramphenicol acetyltransferase (CAT) in the case of pAC184-based hybrid plasmids (*r*). Note that in the case of the pAC184-based hybrid plasmid pMK007, the CAT determinant has been insertionally inactivated by the cloned DNA sequence (*e*). Fragments of polypeptides described in the text are labelled with small solid arrows within the diagram. Plasmid pMK005 and its mutant derivatives were transformed into the *E. coli* minicell strain DS410 (ref. 27). Plasmid-encoded polypeptides were detected by autoradiography and/or fluorography[30] of SDS-polyacrylamide gels of ^{35}S-methionine-labelled minicell preparations[19]. *a*, *f*, pMK005; *b*, pMK006; *c*, pMK002; *d*, pMK001; *e*, pMK007; *g*, pMK053; *h*, pMK062; *i*, pMK063; (*j*) pMK055; *k*, pMK064 (representative of pMK058 and pMK065); *l*, *s*, pMK070 (representatives of pMK067, pMK069, pMK072 and pMK068); *m*, pMK060 (representative of pMK059); *n*, pMK066; *o*, pMK051 (representative of pMK056; pMK057; pMK061 and pMK074); *p*, pMK075 and pMK077; *q*, pMK075; and *r*, pMK077.

pMK063 express additional polypeptides of MWs 21,000 and 19,000 respectively, which seem to be fragments of the K88 antigen resulting from termination of transcription or translation by a signal located within the Tn*5* insertion. In addition, the 17,000-MW band produced by mutant pMK053 is consistently very intense, suggesting that this is a doublet consisting of a ~17,000-MW fragment of the K88 antigen comigrating with the 17,000-MW polypeptide described above. The sizes of these fragments are consistent with the relative positions of the Tn*5* insertions in these mutants.

Mutant pMK055 expressed high levels of the 70,000 and 29,000-MW polypeptides in minicells but failed to produce detectable levels of the 23,500-MW K88 antigen or the 17,000-MW polypeptide (Fig. 3*j*). The Tn*5* insertion in this mutant seems to be located in a cistron (referred to as *adh* C) encoding the 17,000-MW polypeptide. The *adh* C cistron is separated from the *adh* D cistron by a site which accommodates a number of K88$^+$ Adh$^+$ MRHA$^+$ Tn*5* insertions (Fig. 2 and see below). Therefore the absence of the 23,500-MW polypeptide in mutant pMK055 cannot be explained by polarity. It is possible that the 17,000-MW *adh* C product is a positive regulator required for the expression of the *adh* D cistron.

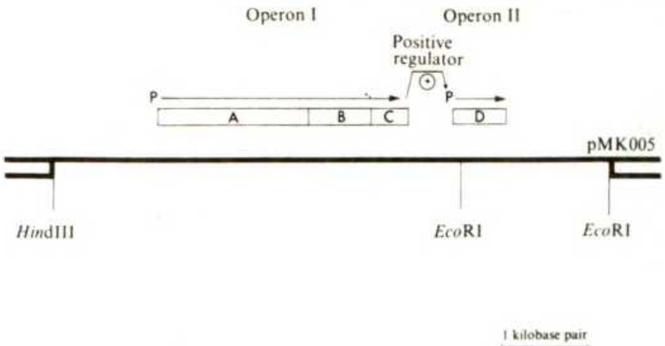


Fig. 5 Model of the genetic organization of the adhesion cistrons. The central solid line represents the cloned DNA sequences in the smallest Adh$^+$ K88$^+$ MRHA$^+$ plasmid, pMK005, the wider shaded boxes at each end represent vector (pBR322) DNA sequences. The boxes above the line represent the *adh* A, *adh* B, *adh* C and *adh* D cistrons; arrowed lines above these indicate the direction of transcription of operons I and II from their respective promoters (P). The 17,000-MW polypeptide product of *adh* C is thought to be a positive regulator required for the expression of operon II (see text).

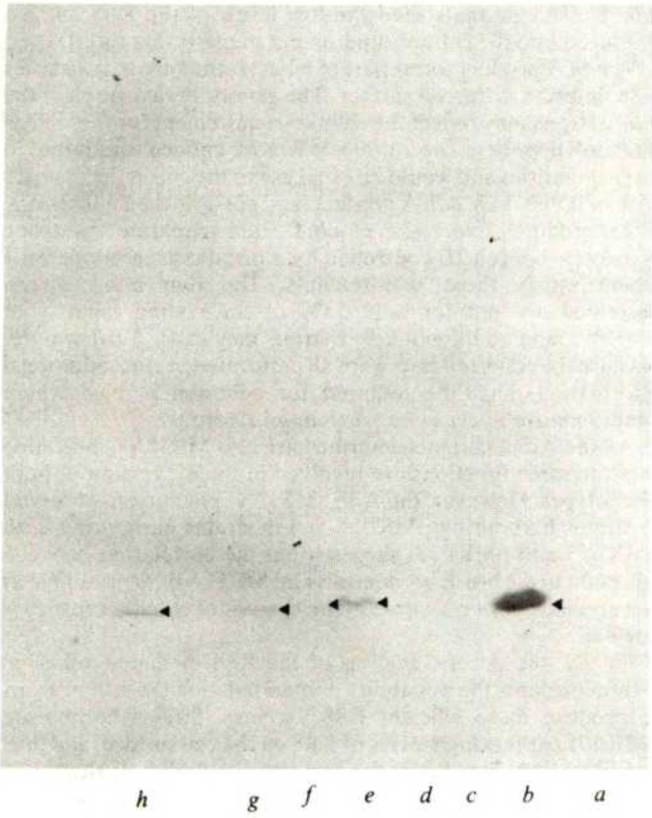


Fig. 4 Immunoprecipitin overlay gel: the method used was a modification of the technique of Broome and Gilbert[31] (B. Crowe, unpublished results). Whole cells harbouring the plasmid being tested were scraped from a nutrient agar plate and resuspended in SDS-polyacrylamide gel final sample buffer, boiled for 2 min and the polypeptides separated on a 12.5% SDS-polyacrylamide gel. When electrophoresis was complete the gel was washed twice (20 min) with gentle shaking in water and soaked for a further 20 min in 15 mM barbitol–HCl buffer, *p*H 8.6, containing 1% (v/v) Triton X-100. The gel was transferred to a glass plate and overlaid with a 1% (w/v) agarose gel in barbitol buffer, containing evenly dispersed anti-K88 antibodies. This was incubated overnight at room temperature in a moist box to allow diffusion of polypeptides from the polyacrylamide gel into the agarose anti-K88 overlay gel. Control experiments using the K88$^+$ Adh$^+$ MRHA$^+$ strain, pMK005, showed that the denatured 23,500-MW K88 polypeptide could be readily detected as a precipitin band in the agarose anti-K88 overlay gel. After incubation the immunoprecipitin agarose gel was washed in 0.1 M NaCl and pressed several times to remove excess antibodies and unprecipitated polypeptides. The immunoprecipitin agarose gel was finally dried in a warm air stream and stained with Coomassie blue. *a*, C600 host; other tracks, C600 harbouring the plasmids *b*, pMK005; *c*, pMK077; *d*, pMK075; *e*, pMK077 and pMK075; *f*, pMK057; *g*, pMK074; *h*, pMK066.

Plasmid pMK066 is the only K88$^+$ Adh$^+$ MRHA$^-$ Tn*5* insertion mutant. Minicells containing this plasmid expressed high levels of the 70,000-MW polypeptide but did not produce the other three polypeptides (Fig. 3*n*). However this mutant produces reduced quantities of the 23,500-MW K88 antigen which was detected by the immunoprecipitin gel overlay technique (Fig. 4*h*). In addition, pMK066 encodes high levels of a new polypeptide of MW 28,000 (Fig. 3*n*). We believe that this polypeptide is a fragment of the 29,000-MW polypeptide which is normally expressed from a cistron (termed *adh* B) whose distal end spans the Tn*5* insertion site in pMK066. This mutation maps adjacent to the left of *adh* C (Fig. 5). The absence of detectable quantities of the 17,000-MW *adh* C product in minicells harbouring pMK066 is explained by suggesting that *adh* B and *adh* C are located in the same operon (operon I, Fig. 5) and that the *adh* B::Tn*5* insertion exerts a strong polar effect on the expression of *adh* C. In the absence of normal levels of the *adh* C product, the level of expression of the *adh* D cistron is reduced. In addition to the 28,000-MW fragment of the *adh* B cistron, pMK066 produces a second new polypeptide of MW 31,000, which can sometimes be resolved from the 30,000 β-lactamase precursor on SDS polyacrylamide gels (data not shown). The origin of this polypeptide is not understood.

The Tn*5* insertions in the remaining K88$^-$ Adh$^-$ MRHA$^-$ pMK005::Tn*5* mutants are located adjacent to each other within a 1.7-kilobase pair region of pMK005 (Fig. 2)—this region is large enough to encode polypeptides totalling a MW of at least 68,000. All these mutants failed to express detectable levels of the four polypeptides in minicells (Fig. 3*m*, *o*). Two mutants (pMK059 and pMK060) which map very close together express a new polypeptide of MW 65,000 which could be a fragment of the 70,000-MW polypeptide. The Tn*5* insertions in all these mutants seem to be located in a cistron (termed *adh* A) which encodes the 70,000-MW polypeptide. The absence of the 29,000 and 17,000-MW polypeptides is explained by suggesting that the *adh* A cistron is located in the promoter proximal part of operon I (Fig. 5) and *adh* A::Tn*5* mutants exert a strong polar effect on the expression of the *adh* B and *adh* C cistrons. The polar effects of *adh* A::Tn*5* mutants and the *adh* B::Tn*5* mutant indicate that operon I is transcribed from left to right as described in Fig. 5. The coding capacity of the 2.8 kilobase pairs of DNA extending from the promoter-proximal *adh*A::Tn*5* insertion, pMK057, to the promoter-distal *adh*C::Tn*5* insertion, pMK055, can be completely accounted for by the *adh*A, *adh*B and *adh*C products (Figs 2, 5).

As the *adh*B::Tn*5* mutant, pMK066, expressed low levels of the 23,000-MW K88 antigen in the absence of normal levels of the *adh*C product the *adh*A::Tn*5* mutants should also express reduced levels of the *adh*D cistron. Representative *adh*A::Tn*5* mutants were analysed by the immunoprecipitin gel overlay technique and were shown to express reduced levels of the K88 antigen (Fig. 4*f*, *g*). However, unlike pMK066, these mutants were Adh$^-$ and were not readily agglutinated by anti-K88 serum in slide agglutination tests (Table 1). Radial immunodiffusion of concentrated cell-free supernatants of these mutants showed that the K88 antigen was present in the culture supernatants (data not shown). These results suggest that in the absence of the 70,000-MW *adh*A product the K88 antigen is shed into the supernatant rather than being assembled into normal K88 fimbriae.

Cloning operons I and II on different compatible plasmids

If the genetic organization of the *adh* cistrons proposed in Fig. 5 is correct it should be possible to obtain high-level expression of the K88 antigen when operons I and II are present in the same host cell but on different compatible plasmids. Furthermore, the 23,500-MW *adh*D product should only be expressed at high levels when complemented in *trans* by operon I, whereas the *adh*A, *adh*B and *adh*C products should be expressed at high levels even when operon I is present alone in the cell. This hypothesis was tested as follows.

The compatible hybrid plasmids pMK075 (pBR322–operon II) and pMK077 (pAC184–operon I) were produced by cloning *Hin*dIII fragments from the pMK005::Tn*5* mutant, pMK070 (Fig. 2). As expected, cells harbouring either pMK075 or pMK077 alone were K88$^-$, Adh$^-$ and MRHA$^-$ (Table 1) and minicells harbouring pMK075 alone did not produce detectable levels of any of the four polypeptides, whereas minicells harbouring pMK077 alone expressed high levels of the 70,000, 29,000 and 17,000-MW polypeptides only (Fig. 3*q*, *r*).

E. coli C600 harbouring both pMK075 and pMK077 expressed reasonably high levels of the K88 antigen (Fig. 4*e*) and were K88$^+$ Adh$^+$ but surprisingly MRHA$^-$ (Table 1). Minicells harbouring both pMK075 and pMK077 expressed high levels of the 70,000, 29,000 and17,000-MW polypeptides but low levels of the 23,500-MW K88 antigen and other pMK075-encoded polypeptides (for example, β-lactamase and its precursors; Fig. 3*p*). It seems that pMK077 and pMK075 do not segregate equally into minicells. However, experiments with whole cells harbouring both of these plasmids (see above) clearly show that pMK075 can express high levels of the 23,500-MW K88 antigen when complemented in *trans* by pMK077. These results support the genetic organization and control of the *adh* cistrons outlined in Fig. 5. However, the MRHA$^-$ phenotype of whole cells harbouring both pMK075 and pMK077 is difficult to understand (see below).

Discussion

Four cistrons, for which we suggested the terms *adh*A, *adh*B, *adh*C and *adh*D, involved in K88-mediated adhesion of enterotoxigenic *E. coli* strains have been identified and mapped (Fig. 5) and their polypeptide products identified in minicells. The 23,500-MW *adh*D product is clearly the K88 fimbrial subunit. The results tentatively suggest the functions of two of the other polypeptides. The 17,000-MW *adh*C product seems to be a positive regulator which controls the expression of *adh*D. However, an alternative explanation to account for the reduction in K88 expression in *adh*C-defective mutants cannot be ruled out. It is possible that in these mutants the *adh*D product has been produced in a precursor form which was not identified by the methods used and that the defect is one of polypeptide processing rather than genetic regulation. The observation that *adh*A::Tn*5* mutants shed the low levels of the K88 antigen produced into the culture supernatant suggests that the 70,000-MW *adh*A product forms part of a basal structure that attaches K88 fimbriae to the cell surface. The genetic organization of the *adh* cistrons may reflect the cellular requirement for the individual polypeptides. The 23,000-MW K88 antigen is required in large quantities and would be of no use to the cell in the absence of the 70,000-MW *adh*A product (and possibly the 29,000-MW *adh*B product). Expression of *adh*D from a separate and strong promoter (operon II) controlled by a distal cistron of operon I would satisfy these requirements. The four *adh* cistrons described account for only 54% of the coding capacity of pMK005 and additional *adh* cistrons may exist. Furthermore, because the adhesion tests were all performed *in vitro*, additional information might be required for adhesion *in vivo* where conditions are likely to be much more rigorous.

All the Adh$^-$ mutants described are also MRHA$^-$, indicating that the same functions are involved in the expression of both phenotypes. However, the Adh$^+$ MRHA$^-$ phenotype, observed in strains harbouring pMK066 and in strains harbouring both pMK075 and pMK077, suggests that the correlation between adhesion to pig brush border cells and MRHA is not as simple as has previously been assumed. The reason for this discrepancy is unclear.

Finally, the genetic analysis of the K88-mediated adhesion system presents the possibility of manipulating the *adh* cistrons to produce more efficient K88 vaccines. Strains harbouring pMK005 express high levels of K88 on the cell surface, and thus would be ideal candidates for oral vaccine studies. It should be possible to construct non-polar *adh*A$^-$ mutants that secrete high levels of the K88 antigen into the culture supernatant and thus facilitate purification of the K88 antigen.

This work was supported by a Wellcome Trust grant to G.D. M.K. is a postdoctoral research assistant employed on this grant. We thank Miss Gillian Johnston and Mrs Corinne Harrison for typing the manuscript and the staff of the preparation room, Department of Microbiology, Trinity College, for supplying sterile glassware and media.

Received 24 November 1980; accepted 5 February 1981.

1. Jones, G. W. in *Receptors and Recognition* Series B, Vol. 3, Ch. 4 (ed. Reissig, J. L.) 141–170 (Chapman and Hall, London, 1970).
2. Arbuthnott, J. P. & Smyth, C. J. in *Adhesion of Microorganisms to Surfaces* (eds Ellwood, D. C., Melling, J. & Rutter, P.) 165–198 (Academic, London, 1979).
3. Mims, C. *Pathogenicity of Infectious Disease* (Academic, London and New York, 1976).
4. Smith, H. W. & Lingwood, M. A. *J. med. Microbiol.* **4,** 467–485 (1971).
5. Smith, H. W. & Lingwood, M. A. *J. med. Microbiol.* **5,** 243–250 (1972).
6. Gurrant, R. L., Moore, P. M., Kirchenfeld, P. M. & Sarde, M. S. *New Engl. J. Med.* **293,** 567–573 (1975).
7. Smith, H. W. *Ciba Fdn Symp.* **42,** 45–72 (1976).
8. Strim, S., Orskov, F., Orskov, I. & Birch-Anderson, A. *J. Bact.* **93,** 340–748 (1967).
9. Evans, D. G., Evans, D. J., Clegg, S. & Pavly, J. A. *Infect. Immunity* **25,** 738–748 (1979).
10. Strim, S., Orskov, F., Orskov, I. & Mansa, B. *J. Bact.* **93,** 731–739 (1976).
11. Nagy, B., Moon, H. W. & Isaacson, R. E. *Infect. Immunity* **16,** 344–352 (1977).
12. Orskov, I., Orskov, F., Smith, H. W. & Sojka, W. J. *Acta path. microbiol. scand.* **83,** 31–36 (1975).
13. Evans, D. G., Silver, R. P., Evans, D. J., Chase, D. G. & Gorbach, S. L. *Infect. Immunity* **12,** 656–667 (1975).
14. Shipley, P., Dougan, G. & Falkow, S. *J. Bact.* (in the press).
15. Shipley, P., Dallas, W. S., Dougan, G. & Falkow, S. in *Microbiology 1979*, 176–180 (AMS, New York, 1979).
16. Bolivar, F. *et al. Gene* **2,** 95–113 (1977).
17. Chang, A. C. Y. & Cohen, S. N. *J. Bact.* **134,** 1141–1156 (1978).
18. Berg, D. E., Weiss, A. & Crossland, L. *J. Bact.* **142,** 436–446 (1980).
19. Dougan, G., Saul, M., Twigg, A., Gill, R. & Sherratt, D. J. *J. Bact.* **138,** 48–54 (1979).
20. Dallas, W. S. & Falkow, S. *Nature* **277,** 406–408 (1979).
21. Rothstein, S. J., Jorgenson, R. A., Postler, K. & Reznikoff, W. S. *Cell* **19,** 795–805 (1980).
22. Mooi, F. R. & de Graaf, F. K. *FEMS Lett.* **5,** 17–20 (1979).
23. Harboe, N. & Ingild, A. *Scand. J. Immun.* **2,** 161–169 (1973).
24. Jones, G. W. & Rutter, J. M. *J. gen. Microbiol.* **84,** 135–144 (1974).
25. Wilson, M. R. & Hohmann, A. W. *Infect. Immunity* **10,** 776–782 (1974).
26. Jorgenson, R. A., Rothstein, S. & Reznikoff, W. S. *Molec. gen. Genet.* **177,** 65–72 (1979).
27. Dougan, G. & Sherratt, D. J. *Molec. gen. Genet.* **151,** 151–160 (1977).
28. Cohen, S. N., Chang, A. C. Y. & Hsu, C. L. *Proc. natn. Acad. Sci. U.S.A.* **69,** 2110–2114 (1972).
29. Meyers, T. A., Sanchez, S., Elwell, L. P. & Falkow, S. *J. Bact.* **127,** 1529–1537 (1976).
30. Bonner, W. M. & Laskey, R. A. *Eur. J. Biochem.* **46,** 83–88 (1974).
31. Broome, S. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **75,** 2746–2749 (1978).

Molecular cloning and characterization of cDNA sequences coding for rat relaxin

P. Hudson*, J. Haley*, M. Cronk†, J. Shine† & H. Niall*

* Howard Florey Institute of Experimental Physiology and Medicine, University of Melbourne, Parkville, Victoria, Australia 3052
† Department of Genetics, Research School of Biological Sciences, Australian National University, Canberra, Australia 2601

Using a synthetic oligonucleotide primer, cloned DNA fragments, each containing the entire coding sequence of rat relaxin, have been isolated from a clone 'bank' of ovarian mRNA sequences. The nucleotide sequence of these clones demonstrates that relaxin is synthesized as a preprorelaxin molecule with an unexpectedly large connecting peptide of 105 amino acid residues.

THE peptide hormone relaxin is produced by the corpora lutea of ovaries during pregnancy in many mammalian species, including pigs[1], rats[2] and humans[3], but also surprisingly in a number of non-mammalian groups (notably the shark)[4]. The secretion of the hormone into the blood stream just before parturition results in a marked softening and lengthening of the pubic symphysis and a softening of the cervix which facilitates the birth process. Relaxin also inhibits uterine contractions and may influence the timing of parturition[5].

The primary structure of porcine relaxin[4,6,7] consists of two peptide chains of 22 and 31 residues, covalently linked by one intra- and two inter-chain disulphide bonds. The locations of the disulphide bonds are homologous to the structures of insulin and insulin-like growth factors (IGFs)[8], for which reason it was anticipated that relaxin is also synthesized as a single-chain precursor polypeptide with a connecting peptide linking the A and B chains. Furthermore, by analogy with the structure of proinsulin[9], it has been suggested that the B chain is at the NH_2-terminus of the precursor molecule.

Nevertheless, porcine relaxin appears to have a precursor of molecular weight (MW) at least 22,000 (22 K) as determined by *in vitro* translation[10], which is much larger than the 11K precursor of insulin, implying either that the connecting peptide of relaxin is far longer than that of insulin (about 30 amino acids) or that there is a peptide extension of the A or B chains which is removed during the processing of the precursor molecule. The recently determined primary structure of rat relaxin[11] has, like porcine relaxin, a limited structural homology with insulin, with an identical disposition of half-cystine residues; the A chain has 24 amino acids, the B chain 35 amino acids.

We report here the *in vitro* synthesis and cloning of cDNA transcribed from total mRNA isolated from pregnant rat ovaries into the bacterial plasmid pBR322. Clones containing the coding sequence for relaxin were detected using a relaxin-specific cDNA probe obtained with the aid of a synthetic oligonucleotide primer. Nucleotide sequence analysis elucidated the entire coding sequence for rat relaxin together with the structure of the 3′-untranslated region, and most of the 5′-untranslated region.

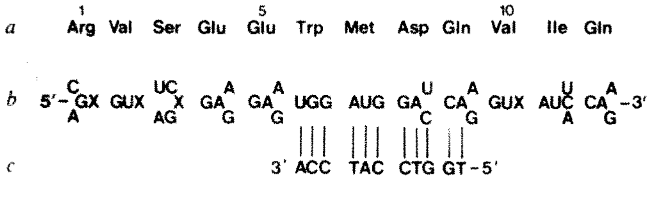

1
5
10
a
Arg Val Ser Glu Glu Trp Met Asp Gln Val Ile Gln
b
5′-GX GUX X GA GA UGG AUG GA CA GUX AUC CA -3′
C A UC A A U A U A
A AG G G C G A G
c
3′ ACC TAC CTG GT-5′

Fig. 1 The choice of an oligonucleotide primer which would specifically hybridize to rat relaxin mRNA. *a*, Amino acids 1–12 of the B chain of rat relaxin. *b*, The hypothetical mRNA sequence coding for rat relaxin, showing ambiguities in the codon choices, where X represents any nucleotide. *c*, The synthetic undecamer finally chosen to be most specific for rat relaxin (see text).

Isolation of cDNA clones

The proportions of relaxin-specific mRNA produced by the corpora lutea of ovaries in pregnant rats are not sufficiently high for the convenient detection and isolation of hormone-specific clones (as with insulin mRNA in the islets of Langerhans[12] and growth hormone in the pituitary[13]). Accordingly, we chose first to generate a cloned cDNA 'bank' from total ovarian mRNA and then sought to detect cDNA clones containing the relaxin-coding sequence using an oligonucleotide primer based on the known amino acid sequence of rat relaxin[11]. (A similar method has been used successfully to detect low levels (0.6%) of gastrin cDNA in a total cDNA population from antral mucosa[14].)

Although the choice of the oligonucleotide primer must be determined from the amino acid sequence of the protein, the degeneracy of the genetic code entails ambiguity at some positions. However, it has been shown[15] that the effects of possible mismatches are minimized by selecting G wherever there is a choice between A and G in the oligonucleotide primer and by selecting T where there is a choice between C and T; the result will either be the correct base pairing or a G–T mismatch. A suitable primer for rat relaxin was chosen by searching through the amino acid sequence[11] to find the region (amino acids 6–9 in the B chain, Trp-Met-Asp-Gln, Fig. 1) for which the nucleotide sequence of the mRNA had the minimum number of ambiguities. We obtained the synthetic primer 5′-TGGTCCATCCA-3′ containing only one possible mismatch within the first 11 nucleotides of this portion of the coding sequence. This primer, being complementary to the mRNA, was used to generate relaxin-specific cDNA from the priming site on the mRNA template (Fig. 2).

mRNA was prepared from 40 ovaries of rats on days 18–20 of pregnancy (term is 21 days) when relaxin levels are highest[16]. Single-stranded cDNA was synthesized using oligo(dT) primer and reverse transcriptase[17] and, after removal of RNA by alkali digestion, converted into double-stranded cDNA by polymerase I (Klenow fragment). This generates a molecule with a hairpin structure on one end and a poly(dA-dT)tract at the other[12]. The hairpin structure was removed by S_1 nuclease and poly(dC) homopolymeric extensions were added onto the 3′ ends with terminal transferase[18].

A clone bank was obtained by inserting the total cDNA population into the *Pst*I endonuclease cleavage site of the bacterial plasmid pBR322[19], using specific hybridization between poly(dG) tails synthesized on the plasmid and poly(dC) tails on the cDNA. An *Escherichia coli* K12 strain (RR1) was transformed with the recombinant molecules (efficiency = 10^5 clones per μg plasmid) and colonies were selected by resistance to tetracycline. Individual colonies were screened on nitrocellulose filters[20] with a radioactive cDNA probe synthesized from the relaxin-specific primer as described above (Fig. 2).

From a total of ~5,000 recombinant clones, 20 colonies hybridized with the cDNA probe (Fig. 2, lane *e*) and also hybridized positively when rescreened with the specific 150-

0028-0836/81/200127—05$01.00

nucleotide fragment eluted from an 8% polyacrylamide gel. Plasmid DNA isolated from 10-ml cultures of each colony was found to contain cloned DNA fragments of 700–900 base pairs. This was not unexpected, because the use of what was essentially a 5′-specific relaxin probe results in the selection of clones carrying nearly full-length cDNA copies of relaxin mRNA.

Extraction of the cDNA insert by cleavage of recombinant plasmids at the *Pst*I sites revealed two types of clone. Sixteen recombinants (type I) contained an internal *Pst*I site in the cDNA insert, giving rise to a 600-base pair DNA fragment (subsequently shown to represent the 3′ portion of the mRNA) and a 150–300-base pair DNA fragment (from the 5′ region of the mRNA). The remaining four recombinants (type II) contained inserts of a similar size but without an internal *Pst*I site.

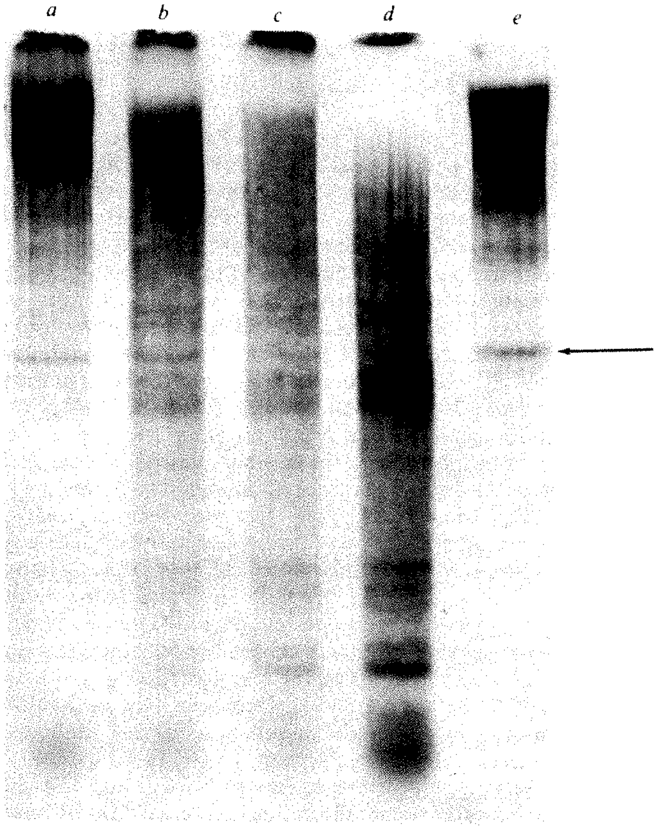


Fig. 2 Primed cDNA synthesis using rat ovarian mRNA and the synthetic undecamer 5′-TGGTCCATCCA-3′ (Collaborative Research Products). mRNA was prepared from rat ovaries, taken on days 18–20 of pregnancy, by extraction in 6 M guanidinium thiocyanate[46] and isolation of polyadenylated RNA on oligo(dT) cellulose[17]. Aliquots (200 ng) of mRNA were transcribed into cDNA in 10-μl reactions containing synthetic primer (1 ng), 50 mM Tris-HCl *p*H 8.0, 50 mM KCl, 7 mM $MgCl_2$, 1 mM 2-mercaptoethanol, 400 μM each of dGTP, dCTP, dTTP, 50 nM $[\alpha\text{-}^{32}P]$dATP (2,000 Ci $mmol^{-1}$), reverse transcriptase (7 units) and various concentrations of nonradioactive dATP (*a*, 1 μM; *b*, 200 nM; *c*, 50 nM; *d*, 0; *e*, 1 μM (preparative scale)) at 37 °C for 20 min. cDNA was denatured in 10 μl of 0.2 M NaOH and heated to 90 °C for 2 min before analysis on a 20 cm × 20 cm 8% polyacrylamide slab gel (150 V, 3 h). The autoradiograph of the dried gel is shown above (Dupont Cronex MRF-31). The predominant cDNA fragment of ~150 nucleotides length (arrowed) is absent in a primed synthesis of non-pregnant rat ovarian mRNA (data not shown) and therefore is probably transcribed from a pregnancy-specific mRNA expected to be rat relaxin. For use in colony hybridizations 10 μg of mRNA were transcribed with 5 ng of primer and a dATP concentration of 1 μM designed to give maximal specific activity in the predominant 150-base cDNA fragment (*e*, above). The reaction mixture was extracted with phenol–chloroform (1 : 1 v/v) in the presence of EDTA (10 mM). Unincorporated nucleotide triphosphates were removed in a column (1 cm × 5 cm) of Sephadex G-50 equilibrated with 0.15 M NaCl, 50 mM Tris-HCl *p*H 7.5 and 1 mM EDTA. cDNA was precipitated with EtOH; dissolved in 0.1 M NaOH and heated to 70 °C for 20 min to remove mRNA. The total cDNA preparation was neutralized and diluted in Denhardt's solution[47] before colony hybridization screens on nitrocellulose filters. Alternatively, the cDNA mixture was fractionated on a 8% polyacrylamide slab gel, the predominant 150-base fragment excised from the gel, electroeluted and precipitated for specific colony screens.

Primary structure of mRNA

The largest plasmid of each type was isolated from 2 l of cells yielding about 1 mg of DNA (pR-RLX1 from type I, pR-RLX2 from type II). The cDNA insert was recovered after *Pst*I digestion and subjected to nucleotide sequence analysis using both chemical and chain termination methods (Fig. 3). The primary structure of relaxin mRNA (Fig. 4) was deduced from the nucleotide sequence of the 840-base pair insert in pR-RLX1. The other type of insert (pR-RLX2) was also completely sequenced, and differed only in the deletion of one nucleotide triplet over the *Pst*I site (GCA or CAG deleted from positions 254–256 or 255–257). The insert within pR-RLX2 was prematurely terminated at nucleotide 77, but was otherwise identical to the structure shown in Fig. 4, thus confirming most of the coding sequence.

The sequence of the 5′-untranslated region does not represent the entire mRNA because the hairpin loop structure of the double-stranded cDNA was removed. However, the size of the specific cDNA synthesized with the relaxin primer (150 nucleotides) suggests that only about 14 nucleotides of the 5′-untranslated sequence are missing from the cloned sequence shown in Fig. 4. It was therefore not surprising to find the clones differing mainly in the size of the 5′ end. A repeat of nucleotides 5–18 over 18–31 may be significant in terms of a regulatory function, but sequences in this region of cloned cDNAs must be viewed cautiously due to the possibility of errors[21] unless confirmed by similar examination of other clones or by direct sequencing of cDNA transcribed from the mRNA[14].

The sequence AAUAAUA in the 3′-untranslated region starting at position 776, about 20 nucleotides from the poly(A) tract, is similar to sequences found in other eukaryotic mRNAs[22,23] and is probably involved in the addition of the 3′-poly(A) sequences after transcription.

The coding sequence of 560 nucleotides (44–604) contains a nonrandom distribution of codon choices, the significance of which is unknown[24].

Amino acid sequence

Figure 4 gives the amino acid sequence for the physiological precursor to rat relaxin (preprorelaxin). The regions corresponding to the known sequences of the A and B chains of rat relaxins are readily identified and the independently derived DNA and peptide sequences for these regions are in complete agreement. The signal peptide of 22 amino acids is followed in order by the B chain (35 residues), a connecting peptide region (105 residues) and the A chain (24 residues). A single stop codon (UGA) follows the cysteine codon for the C-terminal residue of the A chain.

The signal peptide of preprorelaxin contains 22 predominantly hydrophobic amino acids, starting at the initiator codon (AUG, Met) and terminating in Ala before the B chain. There is little homology with the signal peptides of other secreted hormones, apart from the typically high leucine content, which produces some random alignments. A cleavage, presumably by a membrane peptidase after a residue with a small side chain (Ala at position −1), is common of the processing in many signal peptides[25]. The resultant precursor secreted into the endoplasmic reticulum and eventually being packaged into vesicles at the Golgi apparatus is the 164-residue prorelaxin.

The size of the connecting (C) peptide of 105 residues is one of the surprising features of the prorelaxin structure. One of its functions is presumably to direct folding of the precursor so that the correct disulphide bonds are formed between the B and A chains. However, a connecting bridge of about 30 residues is sufficient to perform this function in proinsulin and even smaller bridges of 11 and 8 residues are found in IGF I and IGF II, respectively[26] (Fig. 5). Experiments involving artificial cross-

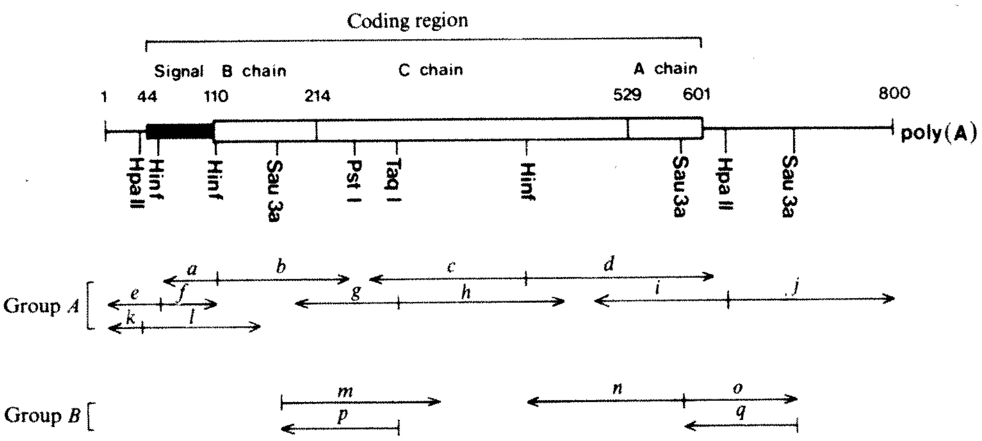


Fig. 3 Schematic diagram of the structural gene for rat relaxin. The nucleotide sequence of 800 base pairs was deduced from the largest recombinant plasmid isolated from the colony screen (pR-RX1, insert of about 900 base pairs into pBR322). Nucleotide residues are numbered in the direction 5′ to 3′ in the message strand, beginning with the first base in the 5′-untranslated region. The poly(dG)·poly(dC) homopolymeric extensions of about 20 nucleotides on each end of the insert are not depicted. The coding region for rat preprorelaxin has been divided into the 66-base coding region corresponding to the signal peptide, the 105-base B chain, 315-base connecting or C peptide and 72-base A chain. The poly(A) extension corresponding to the 3′-terminal sequence of the mRNA structure was found to be about 40 nucleotides long. The insert contained an internal *Pst*I recognition site. Thus, cleavage of plasmid pR-RLX1 with *Pst*I generated a 300-base fragment (containing the 5′-untranslated region) and a 600-base fragment (containing the 3′-untranslated region and poly(A) tract.) The nucleotide sequence of the cloned DNA plasmid was determined using both the method of Maxam and Gilbert[29] and that of Sanger *et al.*[30]. The plasmid was digested with a selection of restriction enzymes which each form 5′-overhanging ends (*Taq*I, *Hin*f, *Hpa*II or *Sau*3a as shown). DNA fragments were end-labelled for the Maxam and Gilbert procedure (group *A*) by incorporating appropriate radionucleotides with reverse transcriptase. Reaction conditions were 50 mM Tris-HCl *p*H 8.0, 50 mM KCl, 7 mM $MgCl_2$, 1 mM 2-mercaptoethanol, [α-^{32}P]dNTP (50 μCi, 2,000 Ci mmol^{-1}) and reverse transcriptase (2 units) in a 20-μl mixture containing the DNA fragment (about 1 μg) at 37 °C for 20 min. After strand separation in alkali, or digestion with a different enzyme, the unique end-labelled fragments were separated on a polyacrylamide slab gel, then detected by autoradiography or by ethidium staining, and electroeluted. The nucleotide sequence was determined by chemical degradation[29] followed by separation of fragments into a sequence ladder on thin 40 cm × 40 cm polyacrylamide slab gels[48]. Fragments sequenced by this procedure are as follows (the letters in italics refer to the horizontal arrows and the labelled ends of the restriction fragments are designated by an asterisk). *a*, *Hin*f–*Hin*f*; *b*, **Hin*f–*Pst*I; *c*, *Pst*I–*Hin*f*; *d*, **Hin*f–*Hpa*II; *e*, *Pst*I–*Hin*f*; *f*, **Hin*f–*Hin*f; *g*, *Sau*3a–*Taq*I*; *h*, **Taq*I–*Sau*3a; *i*, *Hin*f–*Hpa*II*; *j*, **Hpa*II–*Pst*I; *k*, *Pst*I–*Hpa*II*; *l*, **Hpa*II–*Pst*I. Alternatively (group *B*), fragments generated by digestion with *Sau*3a or *Taq*I were subcloned (by ligation to the corresponding complementary ends) into the *Bam*HI or *Acc*I site of bacteriophage M13mp7 and grown after transformation of permissive hosts (*E. coli* strain JMB101). The sequence of these subcloned fragments (*m*–*q*) was determined by primed DNA synthesis using a synthetic primer complementary to the codons for amino acids 13–17 in the β-galactosidase gene. Dideoxynucleotides (P-L Biochemicals) were incorporated using polymerase I (Klenow fragment) in conditions described by Sanger *et al.*[30].

links between insulin A and B chains indicate that a bridge equivalent in length to only one or two amino acids is adequate to direct the correct refolding of previously reduced chains[27]. Thus, the C peptide of prorelaxin vastly exceeds in length the probable minimum requirements for correct alignment of the A and B chains. Although there is no sequence homology between relaxin and insulin C bridges, there is a similar region:

Glu-Asp-Gly-Gly-Pro-Pro-Glu rat relaxin
Glu-Leu-Gly-Gly-Gly-Pro-Glu rat insulin I

with five out of seven identical residues. (Vicinal glycine residues are found in most known insulin C peptides.)

Processing of rat preprorelaxin

As groups of two or more lysine and/or arginine residues are commonly found at sequence regions where processing of peptide hormone precursors occurs[25] and are presumed to act as recognition signals for intracellular proteases, it is surprising that there is no cluster of basic residues immediately adjacent to the C-terminal end of the B chain of relaxin. Not all basic clusters, however, are sites for proteolytic cleavage. For example, the rat relaxin A chain contains the sequence Arg-Arg internally and parathyroid hormone contains two internal basic clusters which are not cleaved in the conversion of proparathyroid hormone to parathyroid hormone[28]. In preprorelaxin there is a cluster of four basic amino acids adjacent to the N-terminus of the A chain (Arg-Lys-Lys-Arg). However, the C-terminal residue of the B chain is leucine 35, and it is followed by the sequence

Ser-Gln-Glu-Glu(AGC CAG GAG GAG)

An examination of the nucleotide sequence in this region (Fig. 4) reveals that a deletion of a single base in the codon for the serine would result in the in-phase sequence AGG AGG, which codes for Arg-Arg. Hence, we examined very carefully the possibility of a sequencing error in this region. However, sequencing experiments using both Maxam-Gilbert[29] and Sanger[30] procedures carried out on both strands of several fragments derived from this region gave unequivocal results in favour of the Ser-Gln-Glu-Glu formulation for the amino acid sequence (see Fig. 3). Furthermore, it can be seen that such a deletion of a nucleotide in this region would lead to a reading frame in which a stop codon appears within 15 nucleotides at position 237–239. A further possibility we considered was an error in the copying of the mRNA sequence by reverse transcriptase. However, this explanation was ruled out because the identical sequence was obtained for this region from another rat relaxin clone (pR-RLX2) (see Fig. 4). This also removes the possibility of a mutation during the cloning procedure itself. A further possibility must be considered, that the processing point is in fact at the next pair of basic residues carboxyl terminal to the present putative cleavage point. This would occur at the Arg-Lys sequence at positions 77–78 of the prorelaxin sequence. In this case, rat relaxin would have a B chain of ~76 residues, and the form which has been isolated from pregnant ovaries and sequenced[11] (which has a B chain of 35 residues) would have to be generated from the larger form by further proteolytic cleavage. We believe this is unlikely because the 6,000-MW (6K) form of relaxin was isolated from ovaries frozen at once in liquid nitrogen and stored at −70 °C until workup by a procedure designed to minimize proteolysis[31]. Further, the 6K form of rat relaxin as isolated by two groups[31,32] is biologically active both in the mouse pubic symphysis bioassay and the rat uterus bioassay. This does not preclude the possibility that there are several forms of rat relaxin generated by variable processing of its precursor. Several high-molecular-weight forms of porcine relaxin have been identified in extracts of pregnant sow ovaries[33]. If the precursor to porcine relaxin is similar to rat preprorelaxin in having a large C-peptide bridge, as seems probable from its molecular weight in *in vitro* translation systems[10], these high-molecular-weight forms of immunologically active relaxin could represent the basic 6K form with variable extensions to A and/or B chains resulting from incomplete removal of the C peptide. Such molecules might represent intermediates in the generation of the biologically active relaxin. However, they might also have a different spectrum of biological properties from the 6K relaxin.

Another possible functional role for the large C peptide may be that it contains peptide sequences with other hormonal activities. This would make prorelaxin analogous to the 31K precursor to ACTH and β-endorphin[34] in containing more than

```
                                                             -20                                     -10
                                                      Met Ser Ser Arg Leu Leu Leu Gln Leu Leu Gly Phe Trp Leu Phe Leu
pR-RLX1 5-CUGAACCGCCCAGGAGCACCGCCCAGGAGCAGCCCAGACCGGA AUG UCC AGC AGA CUC UUG CUC CAG CUC CUG GGG UUC UGG CUG UUC CUG
          1                                                   50

                   -1   1                                                           20
Ser Gln Pro Cys Arg Ala Arg Val Ser Glu Glu Trp Met Asp Gln Val Ile Gln Val Cys Gly Arg Gly Tyr Ala Arg Ala Trp Ile Glu
AGC CAG CCU UGC AGG GCG CGA GUC UCG GAG GAG UGG AUG GAC CAA GUC AUU CAG GUG UGC GGC CGU GGA UAU GCC CGC GCA UGG AUC GAA
         100                                                                150

                                                            40                                  *
Val Cys Gly Ala Ser Val Gly Arg Leu Ala Leu Ser Gln Glu Glu Pro Ala Pro Leu Ala Arg Gln Ala Thr Ala Glu Val Val Pro Ser
GUC UGC GGG GCC UCC GUG GGA AGA CUG GCU UUG AGC CAG GAG GAG CCA GCU CCG CUA GCC AGG CAA GCC ACU GCA GAA GUU GUG CCA UCC
                        200                                                              250

                    60                                                                              80
Phe Ile Asn Lys Asp Ala Glu Pro Phe Asp Met Thr Leu Lys Cys Leu Pro Asn Leu Ser Glu Glu Arg Lys Ala Ala Leu Ser Glu Gly
UUC AUC AAC AAA GAU GCG GAG CCU UUC GAU AUG ACG UUG AAA UGC CUU CCA AAU UUG UCU GAG GAG CGG AAG GCA GCA CUG UCU GAG GGG
                                    300                                                                 350

                                                            100
Arg Ala Pro Phe Pro Glu Leu Gln Gln His Ala Pro Ala Leu Ser Asp Ser Val Val Ser Leu Glu Gly Phe Lys Lys Thr Phe His Asn
CGA GCA CCG UUC CCA GAG CUA CAA CAA CAC GCA CCC GCG UUG AGC GAU UCC GUU GUU AGC UUG GAA GGA UUU AAG AAA ACU UUC CAC AAU
                                                400                                                                 450

                    120                                                                             140
Gln Leu Gly Glu Ala Glu Asp Gly Gly Pro Pro Glu Leu Lys Tyr Leu Gly Ser Asp Ala Gln Ser Arg Lys Lys Arg Gln Ser Gly Ala
CAG CUG GGU GAA GCA GAA GAU GGC GGU CCU CCA GAG CUC AAA UAC UUA GGC UCA GAU GCU CAG UCA CGG AAA AAG AGG CAG UCU GGC GCA
                                                                500

                                                            160
Leu Leu Ser Glu Gln Cys Cys His Ile Gly Cys Thr Arg Arg Ser Ile Ala Lys Leu Cys ***
CUG CUC AGU GAG CAG UGU UGC CAC AUC GGU UGU ACC AGA AGA UCC AUU GCU AAA CUC UGC UGA UACGGAGCUAAUGGUGCGCAACCCGGCAGUGUCG
        550                                                                 600

CGUGCGUUCUUGCUACCUGGCUAAUGCCUCUUUCUGUCCCCUUUUCUUAUUUAUUAGGAUCUGAGAAACCAUAAUUCACACUUGGAUUUGUCCUUUGAUAAGAGAGAAAUGUUCUCUGC
          650                                               700                                               750

AUUGUUGUAGUUCUGUUAAUAAUAUUUUUAUGCAAAAACUUU poly-A ---3'
                                        800
```

Fig. 4 Primary structure of rat preprorelaxin mRNA. The nucleotide sequence of the mRNA was deduced from that of the 900-base pair cDNA insert in pR-RLX1 (compiled in Fig. 3). Nucleotides are numbered sequentially from the 5′ end, being adjacent to the poly(dG) tail in the cDNA. The 5′-terminal sequence is derived solely from one cDNA clone and certainly does not represent the entire sequence of the intact cellular mRNA. However, the 3′ end contains a poly(A) tract of about 40 nucleotides and probably represents the exact mRNA structure. The predicted amino acid sequence shown above is numbered sequentially from the amino terminus of the relaxin B-chain sequence through the 164-residue prorelaxin structure to the termination codon. Therefore, the signal sequence has been numbered from methionine (−22) at the first initiation codon to be found from the 5′ end. The nucleotide sequence of a second clone (pR-RLX2) beginning at position 77 was identical over the coding region to pR-RLX1 with the exception of a deletion of alanine 49 (marked by an asterisk). This could represent the deletion of either nucleotides GCA (254–256) or CAG (255–257).

one hormone within its coding sequence. The presence of two groups of dibasic residues at positions 77–78 and 109–110 of prorelaxin could represent processing points for generation of such putative hormones. At present this suggestion is purely speculative, and based only on the conservation during evolution in at least one and possibly two species of relaxin of what seems to be an unnecessarily large C peptide. There is no information on the fate of the C peptide. However, it should be possible to develop radioimmunoassays using synthetic peptide fragments based on the C-peptide sequence proposed here. Such studies should define whether the C peptide is stored and secreted along with the 6K form of relaxin.

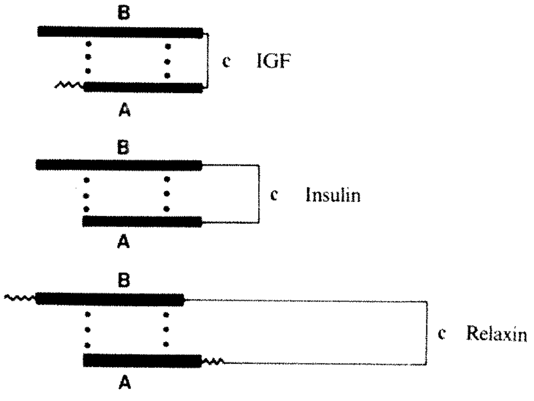


Fig. 5 Schematic comparison of the prohormone structures for insulin-like growth factors[40], insulin[40] and relaxin. These structures represent the precursor form of the hormones after the signal peptide has been removed and before subsequent enzymatic cleavage. The precursor molecules all have structures designated from the amino terminus as B chain: connecting (C) peptide: A chain, with the relative size of the C peptides drawn approximately to scale. The wavy lines represent relaxin and IGF A- and B-chain regions which are nonhomologous to insulin.

The genes for rat preprorelaxin

The finding of two clones encoding slightly different preprorelaxin molecules was fortuitous and resulted from the fact that the single difference (deletion of a base triplet) removed a *Pst*I cleavage site. It seems unlikely that the difference occurred during cDNA synthesis or cloning, as several examples of each type of clone were found. It is more likely that they represent different preprorelaxin mRNAs, whose translation products would differ only by the presence or absence of alanine at position 49 (within the C-peptide region).

There are several possible explanations for the finding of two preprorelaxin mRNAs. As the mRNA was extracted from pooled ovaries from at least 20 pregnant rats from standard laboratory strains not specially selected for genetic uniformity, allelic variation may account for the two mRNAs. Alternatively

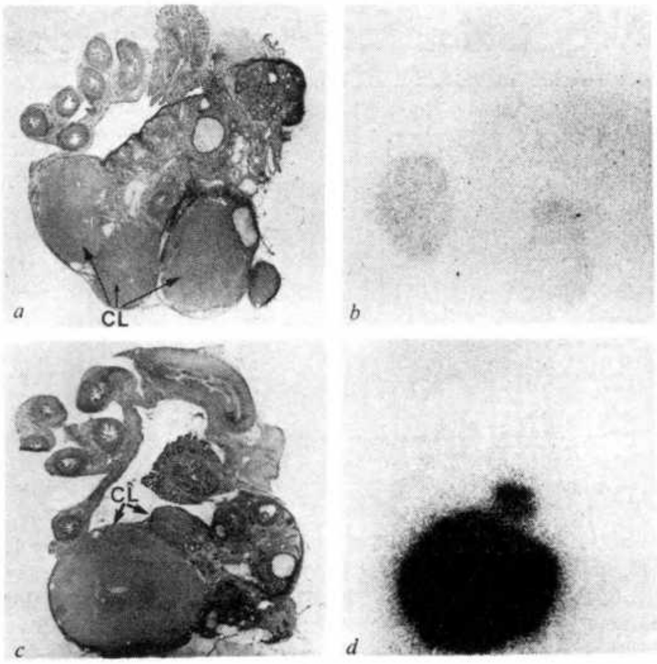


Fig. 6 Hybridization histochemistry[43] used to detect the site of hormone synthesis at the mRNA level in rat ovaries. A relaxin-specific radiolabelled cDNA probe was synthesized from 200 ng of the *Hpa*II fragment of pR-RLX1 (nucleotides 40–628). This represents the entire coding region of rat preprorelaxin with only a small portion of the untranslated regions. The synthesis of the radioactive probe followed the procedure of Taylor *et al.*[49] incorporating [α-^{32}P]dATP with polymerase I (Klenow fragment) and random primers. Tissue sections (6 μm) were cut on a cryostat, fixed and hybridized according to a previous method published from this laboratory[43]. *a*, A non-pregnant rat ovary with corpora lutea marked CL (arrowed) and adjacent stroma, follicles and uterine tube; *b*, autoradiograph of *a* showing background labelling; *c*, a late pregnant rat ovary (18th day) with corpora lutea, adjacent follicles and uterine tube; *d*, autoradiograph of *c* showing heavy labelling of corpora lutea as the site of highest mRNA concentration in the ovary. ×10.

there may be two non-allelic genes for relaxin, as there are for rat insulin[35]. It may be relevant that the deletion found in type II clones is almost exactly in the position of the intervening sequence within the C-peptide regions of human and chicken insulins[36,37] and rat insulin II. This raises a further possibility, that the deletion is the result of a minor variation in excision of an intervening sequence. Such a mechanism has been suggested as a possible source of the 20K version of human growth hormone which differs from the usual 22K form by the deletion of residues 32–46 (refs 38, 39).

Whatever the origin of the two relaxin mRNA species, it seems plausible to suggest that an intron may interrupt the coding sequence for the relaxin C peptide in the genomic DNA. It has previously been proposed that the intervening sequence in the insulin C peptide may account for its being the most highly variable part of the proinsulin molecule, with additions and deletions as well as sequence changes[40]. The striking variability in the size of the C peptide between insulin, relaxin and IGF (Fig. 5) could reflect variations in the relative extent of removal of a common 'genomic C bridge' by successive processing of pre-mRNA and prohormone molecules. These hypotheses can be tested by examining relaxin-encoding clones obtained from cDNA and genomic DNA from individual rats.

Site of hormone biosynthesis in rat ovaries

Immunohistochemical studies have detected relaxin in the corpora lutea of rat ovaries during pregnancy[41] and in cytoplasmic vesicles in pig corpora lutea[42]. These studies support the belief that relaxin is synthesized on ribosomes of the rough endoplasmic reticulum and stored in vesicles of the granulosa lutein cells. We have used an alternative approach, termed hybridization histochemistry[43], which has the advantage of being able to detect in specially prepared whole sections of tissue those areas which contain specific mRNA populations. Using a ^{32}P-labelled cloned cDNA probe corresponding to the coding region of rat preprorelaxin we have identified the major site of mRNA synthesis in ovaries of pregnant rats (Fig. 6). The specificity for the corpora lutea compared with background labelling on adjacent areas of tissue indicates the presence of high levels of relaxin mRNA within the corpora lutea. There was a low level of labelling of the corpora lutea of ovaries from non-pregnant rats due to either nonspecific adsorption or the presence of small amounts of relaxin-specific mRNA in the non-pregnant state. It should be possible to demonstrate the period during which the relaxin gene(s) are expressed by studying the hybridization of the probe to tissue sections taken at various stages throughout pregnancy. Furthermore, when probes corresponding to relaxin mRNA are available from other species such as the human and guinea pig, the technique could be used to examine questions which have been raised[44,45] concerning the site of relaxin synthesis in tissues other than the ovaries.

We thank Jennifer Penschow for preparation of sections for hybridization histochemistry, Dr D. Kemp for terminal transferase and Dr J. Messing for M13mp7. AMV reverse transcriptase was provided by Dr J. Beard through The Office of Program Sources and Logistics, Viral Cancer Program, NCI. This work was supported by grants from the National Health and Medical Research Council of Australia, the Myer Family Trusts, the Ian Potter Foundation, the Howard Florey Biomedical Foundation (USA) and by grant no. HD 11908 (to H.D.N.) from the NIH.

Received 2 January; accepted 12 March 1981.

1. Belt, W. D., Anderson, L. L., Cavazos, L. F. & Melampy, R. M. *Endocrinology* **89,** 1–10 (1971).
2. Anderson, L. L., Bast, J. D. & Melampy, R. M. *J. Endocr.* **59,** 371–372 (1973).
3. O'Byrne, E. *et al. J. clin. Endocr. Metab.* **47,** 1106–1110 (1978).
4. Schwabe, C. *et al. Recent Prog. Horm. Res.* **34,** 123–211 (1978).
5. Porter, D. G. & Downing, S. *J. Reprod. Fert.* **52,** 95–102 (1978).
6. James, R., Niall, H., Kwok, S. & Bryant-Greenwood, G. D. *Nature* **267,** 544–546 (1977).
7. Niall, H. D. *et al.* in *Proc. 15th Mid-West Conf. Endocr. Metab.*, 1980 (ed. Anderson, R.) (Plenum, New York, in the press).
8. Bradshaw, R. A. *Rev. Biochem.* **47,** 191–216 (1978).
9. Tager, H. S. & Steiner, D. F. A. *A. Rev. Biochem.* **43,** 509–538 (1974).
10. Gast, M. J., Mercado-Simmen, R., Niall, H. & Boime, I. *Ann. N.Y. Acad. Sci.* **343,** 148–154 (1980).
11. John, M. J., Walsh, J. R., Borgesson, B. W. & Niall, H. D. *Endocrinology* **108,** 726–729 (1981).
12. Ullrich, A. *et al. Science* **196,** 1313–1319 (1977).
13. Seeburg, P. H., Shine, J., Martial, J. A., Baxter, J. D. & Goodman, H. M. *Nature* **270,** 486–494 (1977).
14. Noyes, B. E., Mevarech, M., Stein, R. & Agarwal, K. L. *Proc. natn. Acad. Sci. U.S.A.* **76,** 1770–1774 (1979).
15. Szostak, J. W. *et al. Meth. Enzym.* **68,** 419–428 (1979).
16. Sherwood, O. D. & Crnekovic, V. E. *Endocrinology* **104,** 893–897 (1979).
17. Bantle, J. A., Maxwell, I. H. & Kahn, W. E. *Analyt. Biochem* **72,** 413–427 (1976).
18. Roychoudhury, R., Jay, E. & Wu, R. *Nucleic Acids Res.* **3,** 863–877 (1976).
19. Chang, A. C. Y. *et al. Nature* **275,** 617–624 (1978).
20. Grunstein, M. & Wallis, J. *Meth. Enzym.* **68,** 379–389 (1979).
21. Richards, R. I., Shine, J., Ullrich, A., Wells, J. R. E. & Goodman, H. M. *Nucleic Acids Res.* **7,** 1137–1146 (1979).
22. Proudfoot, N. J. & Brownlee, G. G. *Nature* **263,** 211–214 (1976).
23. Hobart, P., Crawford, R., Shen, L., Pictet, R. & Rutter, W. J. *Nature* **288,** 137–141 (1980).
24. Grantham, R., Gautier, C., Gouy, M., Mercier, R. & Pave, A. *Nucleic Acids Res.* **8,** 49–62 (1980).
25. Steiner, D. F., Quinn, P. S., Chan, S. J., Marsh, J. & Tager, H. S. *Ann. N.Y. Acad. Sci.* **343,** 1–16 (1980).
26. Rinderknecht, E. & Humbel, R. E. *FEBS Lett.* **89,** 283–286 (1978).
27. Busse, W. D. & Carpenter, F. H. *Biochemistry* **15,** 1649–1657 (1976).
28. Hamilton, J. W. *et al. Proc. natn. Acad. Sci. U.S.A.* **71,** 653–656 (1974).
29. Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74,** 560–564 (1977).
30. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5463–5467 (1977).
31. Walsh, J. R. & Niall, H. D. *Endocrinology* **107,** 1258–1260 (1980).
32. Sherwood, O. D. *Endocrinology* **104,** 886–892 (1979).
33. Kwok, S. C. M., Chamley, W. A. & Bryant-Greenwood, G. D. *Biochem. biophys. Res. Commun.* **82,** 997–1005 (1978).
34. Nakanishi, S. *et al. Nature* **278,** 423–434 (1979).
35. Lomedico, P. *et al. Cell* **18,** 545–558 (1978).
36. Bell, G. I. *et al. Nature* **284,** 26–32 (1980).
37. Perler, F. *et al. Cell* **20,** 555–566 (1980).
38. Fiddes, J. C. *et al. Proc. natn. Acad. Sci. U.S.A.* **76,** 4294–4298 (1979).
39. Wallis, M. *Nature* **284,** 512 (1980).
40. Blundell, T. L. & Humbel, R. E. *Nature* **287,** 781–787 (1980).
41. Anderson, M. L., Long, J. A. & Hayashida, T. *Biol. Reprod.* **13,** 499–504 (1975).
42. Kendall, J. Z., Plopper, C. G. & Bryant-Greenwood, G. D. *Biol. Reprod.* **18,** 94–98 (1974).
43. Hudson, P. *et al. Endocrinology* **108,** 353–356 (1981).
44. Dubois, M. P. & Dacheux, J. L. *Cell Tissue Res.* **187,** 201–214 (1978).
45. Bryant-Greenwood, G. D. in *Proc. 15th Mid-West Conf. Endocr. Metab.*, 1980 (ed. Anderson, R.) (Plenum, New York, in the press).
46. Seeberg, P. H. *et al. Cell* **12,** 157–165 (1977).
47. Denhardt, D. T. *Biochem. biophys. Res. Commun.* **23,** 641–646 (1966).
48. Sanger, F. & Coulson, A. R. *FEBS Lett.* **87,** 107–110 (1978).
49. Taylor, J. M., Illmersee, R. & Summers, J. *Biochim. biophys. Acta* **442,** 324–330 (1976).

LETTERS

Optical spectrum of the unusual supernova remnant G109.1−1.0

William P. Blair & Robert P. Kirshner

Department of Astronomy, University of Michigan, Ann Arbor, Michigan 48109, USA

The new supernova remnant G109.1−1.0 has important X-ray and radio parallels to the SS433–W50 pair. We have obtained a spectrum of the faint optical filaments associated with the large shell-like radio source. The spectrum shows strong [S II] λλ 6,717, 6,731 emission relative to Hα, which is characteristic of shock-heated gas and which confirms that the filaments are part of the supernova remnant, and is generally similar to spectra of the Cygnus Loop. By assuming pressure equilibrium between the optical filaments and the interior of the remnant, we find an initial energy of 3×10^{51} erg, which is higher than that found from X-ray measurements. When compared with W50, the G109.1−1.0 remnant has filaments of higher density, and weaker [N II] emission.

G109.1−1.0 is a newly discovered galactic supernova remnant (SNR) which has been found to be a source of both X-ray[1] and non thermal radio radiation[2]. The X-ray image obtained with the Einstein Observatory shows an incomplete shell of emission with a diameter of ~36 arc min. A strong, compact X-ray source with no known optical counterpart appears at the centre of the X-ray shell. The radio observations show a smaller but complete shell with a spectral index $\alpha\approx-0.5$ ($S_\nu \propto \nu^\alpha$), which is typical for SNRs. These observations also show enhancements in the radio shell along a NE–SW line which approximately line up with a jet-like feature seen both in radio and X ray. This morphology has strong parallels with the association[3–5] between the SNR W50 and the bizarre object SS433.

The distance of G109.9−1.0 has been estimated from its radio surface brightness, Σ, assuming a relationship between Σ and the diameter (D). However, because the angular size of the SNR is different in X-ray and radio maps, and because of differences in the various Σ–D relations, the distance is only roughly known, with estimates ranging from 3.6 to 5.2 kpc. Using the X-ray angular diameter of ~36 arc min, and parameterizing to distance d = 4 kpc, we find the diameter of the SNR is

$$D = 41.9\left(\frac{d}{4\ \text{kpc}}\right)\ \text{pc}$$

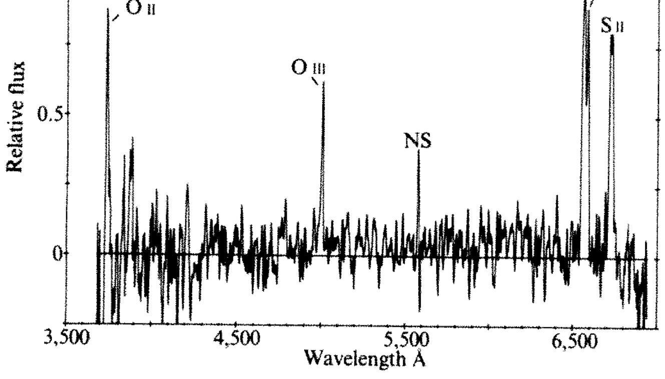


Fig. 1 The optical spectrum of G109.1−1.0 as observed at McGraw-Hill Observatory on 17 October 1980.

Table 1 Line strengths for G109.1−1.0

Line	λ	F(Hα = 100)	I(Hα = 100)†	Cygnus‡ Loop-Pos 2
[O II]	3,727	85	762	430
Hβ	4,861	<15	<48	34
[O III]	4,959	15	44	28
[O III]	5,007	52	151	85
[N II]	6,548	29*	29	29
Hα	6,563	100	100	100
[N II]	6,584	89	86	87
[S II]	6,717	73	66	72
[S II]	6,731	73	66	55

* Assumes [N II] λ 6,584/λ 6,548 = 3.0.
† Assumes $E(B-V)=1.1$.
‡ From ref. 8.

Near the regions of enhanced radio emission are several faint optical filaments visible on the Palomar Observatory Sky Survey (POSS) red plate and on deep Hα and [S II] Schmidt plates[2]. We obtained a spectrum of one of these filaments on 17 October 1980 with the intensified Reticon scanner on the 1.3-m telescope at McGraw-Hill Observatory. This filament is located in the southern portion of the remnant at (1950) α = 22 h 59 min 45 s, δ = 58° 25′ 05″; it runs in a nearly radial direction and is about 2 arc min in length. A 4 × 40 arc s slit was placed along the filament and alternate 10-min integrations were obtained on the object and on a sky position ~10 arc min to the south—well outside the observed emission region. This technique along with the extreme faintness of the filament caused inexact sky subtraction in the resulting spectrum, which is shown reduced to fluxes by standard methods[6] (Fig. 1). The total integration time on the object was 3,000 s.

The observed line fluxes, $F(\lambda)$, relative to $F(\text{H}\alpha)$ = 100 are shown in Table 1. We were not able to detect Hβ in our spectrum and hence cannot determine the reddening directly. However, the reddening has been estimated for two Sharpless HII regions[7] nearby which are thought to be at approximately the same distance as the remnant[1]. Using a median value of $E(B-V)=1.1$, we show these reddening corrected line intensities, $I(\lambda)$, in Table 1. The relative line strengths are accurate to 30% except for [O II] λ 3,727, which is worse because the detector sensitivity is poor and rapidly changing in this region.

We have also listed in Table 1 the relative line strengths[8] of Miller's position 2 in the Cygnus Loop SNR. Comparing these with our observations of G109.1−1.0, we see broad similarities in the relative line strengths. The main difference is that the lines of [O II] λ 3,727 and [O III] λλ 4,959, 5,007 are stronger in G109.1−1.0 although the estimates of these lines are particularly vulnerable to the uncertain reddening. The strong lines of [S II] λλ 6,717, 6,731 and [N II] λ 6,584 relative to Hα are the usual indicators of shock heating[9] and this confirms that we are seeing the optical filaments of a SNR.

Our spectrum is not good enough to warrant a detailed analysis and comparison with published shock models, but we can use the relative strengths of the [S II] λλ 6,717, 6,731 lines to estimate the electron density in the S^+ zone of the emitting region. Using updated collisional cross-sections for [S II][10] and a five-level atom calculation, we estimate N_e(S II) ≈700 cm^{-2}.

As the remnant is not bright optically, it is probably still in the adiabatic expansion phase. The optical filaments are probably dense clouds which have been overtaken by the blast wave from the supernova as it travels through the interstellar gas. Using the formulation of McKee and Cowie[11], we can estimate the initial energy of the supernova by assuming rough pressure equilibrium between the clouds and intercloud gas (now emitting X rays), given by

$$E_0 = 5.6\times10^{43}(\beta')^{-1}N_e(\text{S II})D(\text{pc})^3\ \text{erg}$$

0028-0836/81/200132—02$01.00

here β' is a factor of the order of unity which relates the pressure in the cloud and intercloud gas. Using values of N_e(S II) and D mentioned earler, we find

$$E_0 = 2.88\times 10^{51}\left(\frac{d}{4\ \mathrm{kpc}}\right)^3\ \mathrm{erg}$$

This is considerably greater than the value of $\sim 7\times 10^{50}$ erg which has been found for the Cygnus Loop (which is about the same diameter) using the same method[6]. More importantly, it is more than an order of magnitude greater than the initial energy calculated from the X-ray luminosity. Using the X-ray parameters derived by Gregory and Fahlman[1], we find

$$E_0(\mathrm{X\ ray}) = 1.84\times 10^{50}\left(\frac{d}{4\ \mathrm{kpc}}\right)^3\ \mathrm{erg}$$

This type of discrepancy for SNR has been noticed previously[12] and may indicate either that the magnetic pressure in the filaments is non-negligible or that there are beams from the central object, as in W50.

The presence of a central radio and X-ray source in G109.1−1.0 and a possible jet-like structure[1] indicate a striking similarity to W50. The emission-line object SS433, which is believed to be associated with W50, also shows a double jet structure in both X rays[3] and radio[13]. The axis of this jet seems to be aligned with the enhanced 'radio ear' structure of W50 and it is in this region where faint optical filaments are visible on the POSS red print. Spectra of some of these filaments[5,14,15] indicate a qualitative similarity to the position we have observed in G109.1−1.0, except that in W50 the [N II] and [S II] lines are much stronger relative to Hα and the density as measured from the [S II] lines is lower in W50.

R.P.K. is an Alfred P. Sloan research fellow whose research on supernovae is supported by NSF grant AST77-17600.

Received 17 February; accepted 16 March 1981.

1. Gregory, P. C. & Fahlman, G. G. *Nature* **287,** 805–806 (1980).
2. Hughes, V. A., Harten, R. H. & van den Bergh, S. *Astrophys. J.* (in the press).
3. Seward, F., Grindlay, J., Seaquist, E. Gilmore, W. *Nature* **287,** 807–809 (1980).
4. Geldzahler, B., Pauls, T. & Salter, C. *Astr. Astrophys.* **84,** 237 (1980).
5. Kirshner, R. P. & Chevalier, R. A. *Astrophys. J. Lett.* **242,** L77–L81 (1980).
6. Blair, W. P., Kirshner, R. P., Gull, T. R., Sawyer, D. L. & Parker, R. A. R. *Astrophys. J.* **242,** 592–595 (1980).
7. Sharpless, S. *Astrophys. J. Suppl.* **41,** 257–279 (1959).
8. Miller, J. S. *Astrophys. J.* **189,** 239–248 (1974).
9. D'Odorico, S. & Sabbadin, F. *Astr. Astrophys.* **53,** 443–445 (1976).
10. Pradhan, A. K. *Mon. Not. R. Astr. Soc.* **183,** 89P–92P (1978).
11. McKee, C. F. & Cowie, L. L. *Astrophys. J.* **195,** 715–725 (1975).
12. Blair, W. P., Kirshner, R. P. & Chevalier, R. A. *Astrophys. J.* (in the press).
13. Spencer, R. E. *IAU Circ.* No. 3521 (1980).
14. Murdin, P. & Clark, D. H. *Mon. Not. R. Astr. Soc.* **190,** 65P (1980).
15. Shuder, J. M., Hatfield, B. F. & Cohen, R. D. *Publ. astr. Soc. Pac.* **92,** 259 (1980).

Cosmological consequences of grand unified theories on density fluctuations

David Lindley

Astronomy Centre, University of Sussex, Falmer,
Brighton BN1 9QH, UK

Recent investigations into the cosmological consequences of grand unified theories (GUTs) of elementary particles have shown that the observed matter–antimatter asymmetry of the Universe can be explained without recourse to the hypothesis of specific initial conditions. It is shown here that the origin of inhomogeneities in the matter distribution, which are thought to be responsible for the later formation of galaxies, cannot be explained by a simple addition of density fluctuations to the standard model. The appearance of these fluctuations, after the epoch when baryon number is fixed, is almost purely adiabatic, any departure from adiabaticity falling off in inverse proportion to the mass of the perturbation.

Fry *et al.*[1] have calculated detailed models for the creation of baryon number in the early stages of the Universe. Their results are parametrized by a quantity $K = 2.9\times 10^{17}\alpha(m_X/\mathrm{GeV})^{-1}$, where $\alpha \simeq 10^{-2}$ is the GUT coupling constant, and m_X is the mass of the superheavy X-boson whose decay is responsible for the generation of baryon number. If $K \ll 1$, the X-bosons freeze out of thermal equilibrium while they still have relativistic energies, and their number density relative to other particles is independent of the freeze-out temperature and of K. If, on the other hand, $K > 1$, the freeze-out temperature is less than m_X, and the number of X-bosons is reduced by a Boltzmann factor. The resulting ratio B of baryon to photon number then decreases with increasing K.

In models of slightly inhomogeneous cosmologies, one conventionally imagines the Universe to consist of a patchwork of pieces, each larger than the horizon at any given time, and each evolving like a separate Friedmann model with a particular value of curvature[2]. In one of the perturbed regions, the expansion rate will be slightly altered by the curvature term, and this will mean that the freeze-out epoch t_f and temperature T_f will differ from the corresponding values in the unperturbed region (which is taken to have zero curvature). If $K < 1$, this variation in T_f will have no effect on the final value of B, but if $K > 1$, B depends on T_f through the Boltzmann factor. (Strictly speaking, the dependence of B on T_f only vanishes in the limit $T_f/m_X \to \infty$, but it falls off rapidly as the X-bosons become relativistic at freeze-out.) Consider a perturbed region just entering the horizon at t_f. If the density variation is $\delta\rho/\rho = \kappa$, the expansion rate[2] will be

$$(\dot{R}/R) = \left(\frac{8\pi G}{3}\rho\right)^{1/2}(1-\kappa)^{1/2} \simeq \left(\frac{8\pi G}{3}\rho\right)^{1/2}\left(1-\frac{\kappa}{2}\right) \qquad (1)$$

Fry *et al.*[1] give a simple analytic expression for B which applies in the limit $K \gg 1$. Using a variable $z = m_X/T$, then $B = \varepsilon/z_f K$, where ε is the baryon number produced in the decay of an $X\bar{X}$ pair; z_f denotes the freeze-out temperature, and is the larger of the two solutions of

$$Kz^{7/2}\,\mathrm{e}^{-z} = 1-\frac{\kappa}{2} \qquad (2)$$

This equation, although like that given by Fry *et al.*, incorporates the modified expansion rate of equation (1). Now, let z_f be the solution of equation (2) in the unperturbed ($\kappa = 0$) region, and let $z_f' = z_f + \mathrm{d}z_f$ be the solution for non-zero κ. It is easy to show that $\mathrm{d}z_f/z_f \simeq \kappa/(2z_f - 7)$. The bad behaviour at $z_f \simeq 7/2$ is because equation (2) is an approximation which fails for small z; we therefore require $z_f \gg 7/2$, and take $\mathrm{d}z_f/z_f \simeq \kappa/2z_f$. With the expression for B, then $\delta B/B = -\mathrm{d}z_f/z_f \simeq \kappa/2z_f$. The minus sign means that regions of higher density (positive κ) have smaller values of baryon-to-photon ratio, because freeze-out occurs later in the overdense regions, where the expansion rate is reduced. These conclusions were reached qualitatively by Turner and Schramm[3] and by Barrow[4].

It can be concluded that if K is small, density fluctuations do not alter B: that is, they are entirely adiabatic. If $K > 1$, then departures from adiabaticity can be achieved, but the effect decreases with increasing K (which implies increasing z_f). There must therefore be some pair of values of K and z_f (presumably near unity) for which the non-adiabaticity is maximized. The appearance of the fluctuations is, in a sense, the opposite of isothermal, because the variation in baryon density is smaller than that in radiation density. Note that these departures from adiabaticity have no dynamical significance at these early stages of cosmic evolution, because at these times baryon number is simply a quantum number attached to the various particles comprising the radiation gas. Variations in baryon density reflect variations in the locally averaged baryon number at different points. Only at a later stage ($t \simeq 10^{-4}$ s), when nucleons become massive (non-relativistic) particles carrying the baryon number, does the distinction between adiabatic and isothermal fluctuations have a dynamical role.

0028-0836/81/200133—02$01.00

The result derived above applies directly only to those fluctuations entering the horizon at t_f. A region much larger than the horizon at this time will have the same relation betwen κ, $\delta\rho/\rho$ and $\delta B/B$ at t_f, but its subsequent evolution must be taken into account. The growth law for large perturbations is $\delta\rho/\rho \propto t$. Bardeen[5] has shown that the time-dependence of perturbed quantities depends on their definition. The analysis here, using different Friedmann models for different regions, corresponds, in Bardeen's terminology, to a description based on co-moving spatial hypersurfaces, for which the quoted growth law applies in a radiation dominated universe. By the time a large perturbation comes within the horizon, at t_h, its density contrast will have been increased by the factor t_h/t_f. Meanwhile, however, the values of B in perturbed and unperturbed regions will have remained constant, because they are quantities purely local to each region, and so $\delta B/B$ will have also remained constant. Thus the departure from adiabaticity will be reduced by a factor t_h/t_f, and as the horizon mass is proportional to t, the non-adiabatic part of the perturbation falls off in inverse proportion to the mass of the inhomogeneity. Because $t_f \simeq 10^{-33}$ s, and t_h can be as large as 10^{12} s for a region entering the horizon at recombination, the character of perturbations large enough to be of interest for galaxy formation will be purely adiabatic to a very high degree.

Hogan[6] has recently shown that, in a hot big bang, cosmological fluctuations induced by a phase transition at a later stage of cosmic history than is considered here are inevitably insufficient in magnitude to account for galaxy formation. In a 'tepid' big bang, such spontaneous fluctuations may be strong enough, but this requires $B \geq 1$ in the early stages[6], which cannot be provided by GUTs[7]. I have shown here that density fluctuations built into the initial conditions of the Universe are also incapable of being directly responsible for large matter inhomogeneities. A possible solution to this dilemma may lie in the generation of inhomogeneities through some kind of non-linear phenomenon (such as the 'Purcell effect'[8], although this requires strong adiabatic fluctuations, $\delta\rho/\rho \simeq 10\%$), or perhaps through the use of evaporating primordial black holes as a source of baryons[3].

Barrow and Turner[9] have performed a similar calculation to that presented here, but suggest that anisotropic (shear) fluctuations could produce useful isothermal variations because they alter the expansion rate at early times, but, unlike the modes considered here, subsequently decay rather than grow. Although the assumed presence of shear inhomogeneities is reasonable, at the same time one must ensure that density fluctuations are absent, to better than ~1 part in 10^{40} of the anisotropic modes, otherwise they will eventually grow and dominate. It thus seems that rather special initial conditions would be needed to create significant isothermal fluctuations.

I thank Professor R. J. Tayler for advice on the presentation and the SRC for a research studentship. I also thank John Barrow for a preprint.

Received 16 January; accepted 6 March 1981.

1. Fry, J. N., Olive, K. A. & Turner, M. S. *Phys. Rev.* **D22**, 2953–2977 (1980).
2. Harrison, E. R. *Phys. Rev.* **D1**, 2726–2730 (1970).
3. Turner, M. S. & Schramm, D. N. *Nature* **279**, 303–305 (1979).
4. Barrow, J. D. *Surv. High. Energy Phys.* **1**, 3 (1980).
5. Bardeen, J. M. *Phys. Rev.* **D22**, 1882–1905 (1980).
6. Hogan, C. J. *Nature* **286**, 360–362 (1980).
7. Nanopoulos, D. V. & Weinberg, S. *Phys. Rev.* **D20**, 2484–2495 (1979).
8. Press, W. H. & Vishniac, E. T. *Nature* **279**, 137–139 (1979).
9. Barrow, J. D. & Turner, M. S. Preprint (University of California, Santa Barbara, 1981).

Meteoroid impact into short-period comet nuclei

Mark J. Cintala

Department of Geological Sciences, Brown University, Providence,
Rhode Island 02912, USA

After their introduction into the inner Solar System, most short-period comets spend the majority of each orbit within the asteroid belt. Provided that cometary material is similar to terrestrial snow in its shock behaviour, the nuclei would accumulate meteoroids as a consequence of collisions with asteroidal debris. Here the effects which extended meteoroid bombardment might have on short-period comet nuclei in ideal conditions are evaluated; the results are extrapolated to more probable situations, and their possible implications for asteroid and meteorite studies are considered. It is suggested that after devolatilization of the nucleus, the accumulated material might combine with primary cometary solids to produce environments conducive to the formation of polymict brecciated meteorites and to influence reflectance spectra of the 'asteroid'. In addition, such meteoroids removed from active nuclei during degassing could be placed in trajectories which might evolve into Earth-crossing orbits.

It is assumed here that comet nuclei are H_2O-rich bodies with weak, probably porous structures, containing a non-negligible fraction of non-volatile 'dust'[1]. Deduction of various physical properties suggests that the impact behaviour of nuclear material can be approximated by that of terrestrial snow (Table 1). Experimental studies reveal that a projectile impacting slightly packed snow will penetrate to many multiples of its own length[2]; its final location is generally at the apex of a conical 'crater'[2]; and the 'crater' is characterized by a large depth/diameter ratio, apparently due to a large amount of target compression[3]. A least-squares fit to data by Swinzow[2] for the penetration (L) of 0.55-cm metal cubes into snow (0.41–0.42 g cm^{-3}) yields

$$L = 0.2\, d\rho^{1.045} V_i^{0.349} \qquad (1)$$

where d, ρ and V_i are the projectile length, density and impact velocity, respectively (all variables in c.g.s. units). While some of the experiments used in deriving equation (1) involved impact velocities lower ($\leqslant 1.3$ km s^{-1}) than those considered to be ideal, many were conducted in conditions expected to apply for the cases treated below—within or near the snow–ice transition regime[4,5]. As spallation will remove mass from silicate projectiles during their penetration at high velocities, equation (1) should yield upper limits to penetration depths.

Assuming that a comet is perturbed into an orbit lying completely within the asteroid belt, then the total mass flux (ϕ) of particles impacting a non-gravitating surface of unit area can be written as

$$\phi = \frac{k\alpha}{\alpha + 1}\left(m_0^{\alpha+1} - M^{\alpha+1}\right) \qquad (2)$$

where k and α are constants, m_0 is the minimum particle mass

Table 1 Brief comparison of the physical properties of H_2O snow and those deduced for comet nuclei

Property	H_2O snow Value	Ref.	Comet nuclei Value	Ref.
Density	0.1–0.9 g cm^{-3}	22	−0.2–1.3* g cm^{3}	7, 23*
Compressive strength	10^4–5×10^7 dyn cm^{-2} †	5	10^4–10^6 dyn cm^{-2}	24
Tensile strength	10^4–10^7 dyn cm^{-2} †	5	~10^3 dyn cm^{-2}	25

* Includes nonvolatile solids. †Over a density range of 0.1–0.9 g cm^{-3}.

0028-0836/81/200134—03$01.00

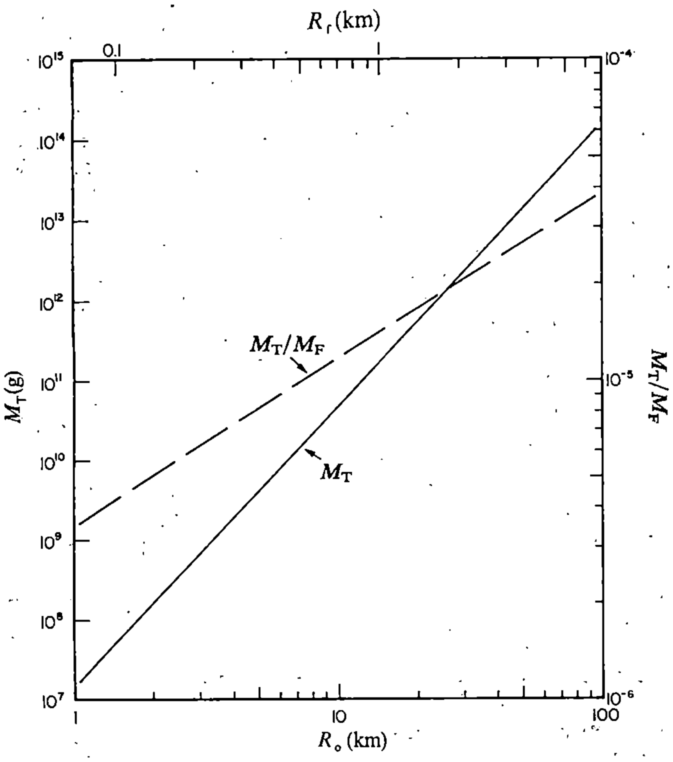


Fig. 1 Total mass of debris (M_T) accumulated by comet nuclei of various initial radii (R_0). M_F is the mass of the burned-out nucleus, exclusive of the accumulated debris, and R_f is the final radius of the devolatilized nucleus. For comparison, the lunar mare basalts comprise between 4×10^{-5} and 4×10^{-4} times the total lunar mass.

that can be retained against gas drag during the nucleus' devolatilization, and M is the largest mass which could be stopped within the average chord length of the (assumed spherical) nucleus. Projectiles with masses greater than M will either disrupt or completely penetrate the nucleus, the former of which would nullify the applicability of the following calculations. Using the equation derived by Whipple[1] to fix m_0 and equation (1) to establish M results in

$$\phi(t)=\frac{k\alpha}{\alpha+1}\left[\left(\frac{2.87\times10^{19}\rho}{nq^{27/4}R_c^3(t)}\right)^{\alpha+1} - (253.67R_c^3(t)\rho^{-2.135}V_i^{-1.047})^{\alpha+1}\right] \quad (3)$$

where $1/n$ is the fraction of incident solar radiation used in subliming ices and $R_c(t)$ is the time-dependent nuclear radius. The perihelion (q) is expressed in astronomical units. Taking the rate of change of the nuclear radius (λ) to be constant, its surface area at time t can be written as

$$A(t)=4\pi R_c^2(t)=4\pi(R_0-\lambda t)^2 \quad (4)$$

where R_0 is the initial radius. The total mass acquired by the nucleus from time 0 (introduction to the asteroid belt) to time T (total devolatilization) can be found from

$$M_T=\int_0^T \phi(t)A(t)\,dt \quad (5)$$

Only silicate projectiles will be treated here, making a chondritic density of 3.5 g cm^{-3} appropriate, and the average asteroidal encounter velocity of 5 km s^{-1} (ref. 6) will be used for this case. Conservative values of 2.0 AU and 1 will be assumed for q and n, respectively. Estimates made by Sekanina[7] yield a value of 2.84×10^{-6} cm s^{-1} for λ and a mean final radius (R_f) of $0.069\ R_0$. The impact flux in the asteroid belt suggested by Housen *et al.*[8] will be adopted, setting k and α at 4.74×10^{-17} and -0.833, respectively. Along with the fact that $R_f=R_0-\lambda T$, inserting these quantities in equation (5) yields

$$M_T=3.5\times10^{-10}(0.139R_0^{3.501}-1.204\times10^3R_0^{2.499})\ \text{g}. \quad (6)$$

The results of equation (6) for nuclei of various initial radii are shown in Fig. 1 which also includes the total accumulated masses normalized to the final nuclear mass (assuming $R_f=0.069\ R_0$ and a density of 3.0 g cm^{-3}).

It is anticipated that these former projectiles would, after devolatilization of the nucleus, be concentrated in the near-surface layers of the 'regolith'. The surface density of this 'foreign' debris (assuming an even distribution over the surface) can be found by dividing equation (6) by the final surface area of the nucleus ($A_f=4\pi R_f^2$)

$$\mu=5.85\times10^{-9}(0.139R_0^{1.501}-1.204\times10^3R_0^{0.499})\ \text{g cm}^{-2} \quad (7)$$

which, in turn, can be translated into an effective thickness (h) by

$$h(p)=\frac{\mu}{\rho(1-p)} \quad (8)$$

where p is the porosity of the surface layer. Values of $h(p)$ are plotted in Fig. 2 for various porosities and initial nuclear radii.

The orbit used in the calculations is not typical of most short-period comets[9], as most observed orbits are significantly more eccentric than that assumed here. Comets with highly inclined orbits are excluded from discussion, as they would spend negligible time in the plane of the asteroid belt. The more eccentric orbits will increase the average impact velocity of asteroid debris. Equation (3) shows, however, that the effects on $\phi(t)$ will be minor: a difference in V_i of a factor of ±5 will change $\phi(t)$ by a factor of only ∓0.3. The time spent outside the belt (hence, in a less frenzied impact environment) will also depend on the orbital eccentricity. Nonetheless, most short-period comets with aphelia inside the orbit of Jupiter spend a large fraction ($\geq75\%$) of each orbital period within the asteroid belt[10], thus placing minimal restrictions on the model results. Finally, the values of λ and $1/n$ used for the model orbit are almost certainly too high. At an orbital average distance of 2.8 AU, for example, water snow will sublime at $\sim9.2\times10^{-7}$ cm s^{-1} (ref. 11)—a rate lower than that used above by more than a factor of 3. Considering the fact that the average perihelion distance of the objects used in deriving λ is ~1.3 AU, the results given by equation (6) should be more appropriate for the

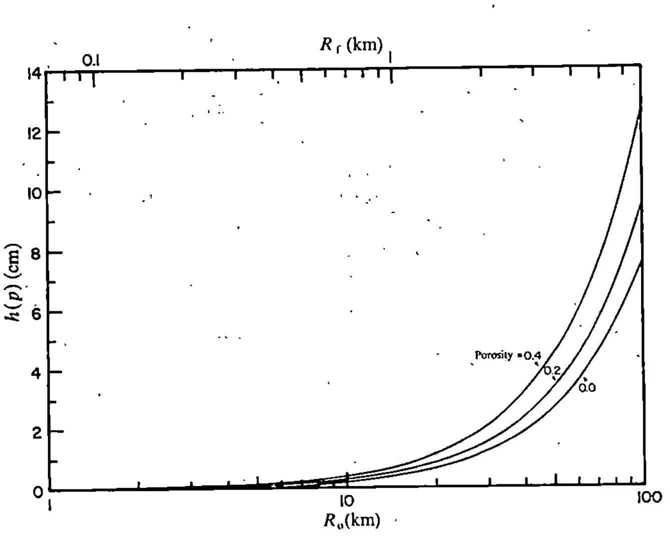


Fig. 2 Effective thickness of the accumulated debris layer as a function of porosity [$h(p)$] compared with the initial and final radii of the nucleus (R_0 and R_f, respectively). Most natural fragmental deposits have porosities of ~0.4. Projectile density: 3.5 g cm^{-3}.

Effective surface layer density (g cm^{-3}):	3.5	2.8	2.1
Porosity:	0	0.2	0.4

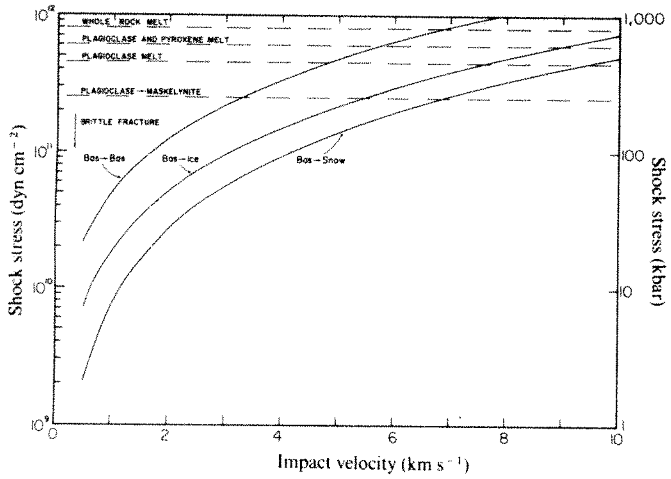


Fig. 3 One-dimensional shock stress as a function of impact velocity for basalt projectiles incident on basalt, H_2O ice, and Greenland snow. Dashed horizontal lines mark the stresses at which the indicated phase changes were observed to occur in experimentally shocked basalts[21].

average rate of radius decrease for more eccentric orbits. While ejection and vaporization of volatile and non-volatile material during impact events would also contribute to mass loss from the nucleus, its magnitude relative to the solar irradiation-induced component is unknown. The value of 1 used for $1/n$ implies that the albedo of the nucleus is effectively zero, thus yielding the unlikely condition that all incident radiation is absorbed and used in subliming the volatile ices. Clearly, this is not the case; a smaller average fraction of solar radiation will be partitioned into devolatilization of the nucleus. These values for λ and $1/n$ will yield a shorter lifetime for a given nucleus than would actually be the case. The calculated masses and effective thicknesses of debris should therefore be treated as conservative estimates *vis à vis* these parameters.

According to recent evidence[12], polymict brecciated meteorites might have been formed during low-velocity collisions between asteroids of dissimilar compositions. Another mechanism which could also account for some of these meteorite types might be the shock lithification of the variegated regolith of a burned-out short-period comet nucleus, such as that described above. Stresses sufficiently high to indurate particulate regoliths would be generated at the impact velocities expected to occur in the asteroid belt[10]. Thus, the generation of polymict breccias on such 'foster-parent' bodies would be directly analogous to the formation of their lunar counterparts. Insofar as the foreign fragments were, in all probability, originally components of asteroid regoliths[8,13], the observed abundance of implanted solar wind gases and solar flare/cosmic ray tracks should be consistent with this model.

Such a regolith would present a reflectance spectrum characteristic of the averaged behaviour of its component lithologies and/or mineralogies. While often considered to be carbonaceous, the composition of primary cometary debris is uncertain[13,14]. Thus, the spectral signature of a devolatilized nucleus cannot be predicted with confidence, suface exposures of foreign debris notwithstanding. For this reason, it would be a formidable task to select candidate objects on the basis of spectrophotometric grounds alone. Note that observations would be biased towards the larger objects, implying that the original (volatile-rich) nuclei would also have been large relative to most presently-active short-period comets[15]. This model predicts that more massive nuclei would have collected more debris, thus maximizing any spectral effects resulting from the overall process. 'Mixing models' similar to those applied by Chapman and Gaffey[16] might yield information relating to the efficiency of this mechanism. Orbital and dynamical studies of fireball and meteorite orbits present strong arguments in favour of some relationships between meteoroids and short-period comets[17,18]. Objections presented by petrologic and chemical arguments against cometary origins of non-carbonaceous meteorites[13,19] might be alleviated if the comets acquired the meteoroids during passage through the asteroid belt or at some earlier time in their histories. Figure 3 illustrates that no petrographically observable shock effects other than brittle fracture would occur in basalt-like assemblages as a result of collisions with comet nuclei at velocities lower than ~5–7 km s⁻¹. Comets in highly eccentric orbits (which characterize Earth-crossing comets, such as P/Encke) would encounter asteroidal debris at higher velocities (D. R. Davies and E. M. Shoemaker, personal communications), forcing the projectiles to incur greater shock damage. Fragments 'absorbed' by short-period comets in less elliptical orbits, however, would be subjected to relatively mild stress levels. Ejection of these meteoroids during periods of devolatilization would place them in orbits similar to that of the nucleus. If the nucleus' orbit were conducive to Jupiter-related perturbations[16], or if it had become Earth-crossing (for example, due to perturbations by Jupiter, or by nongravitational forces[20]), the ejected fragments would then be in orbits which could ultimately encounter the Earth. The total annual mass which might be delivered to the Earth by this mechanism is uncertain, but large fragments (of the order of 10^4–10^6 g, depending on the nuclear radius, perihelion, and so on[1]) can be removed from nuclei by gas drag. Because of the ease with which they lose solid debris, small nuclei should be more effective protagonists in this process.

Finally, a spacecraft with high-resolution (≤ 1 m) imaging, multispectral, and/or chemical mapping capability could survey the surface exposure of material on an active short-period comet nucleus. Such an investigation would provide important data regarding the composition, abundance and dispersal of debris in the regions of the asteroid belt traversed by the comet during its active lifetime.

This work was carried out under NASA grant NGR-40-002-088. Helpful reviews, comments and/or discussions were provided by A. Brown, J. Head, P. Mouginis-Mark, J. Whitford-Stark, L. Wilson, C. Chapman, D. Davis, F. Fanale, E. Shoemaker, J. Veverka, F. Whipple and B. Zellner. Many of the ideas presented here were conceived during conversations with members of the US Army Cold Regions Research and Engineering Laboratory in Hanover, NH. Nancy Christy and Sue Sharpton assisted with manuscript preparation.

Received 5 September 1980; accepted 27 February 1981.

1. Whipple, F. L. *Astrophys. J.* **111**, 375 (1950); **113**, 464 (1951).
2. Swinzow, G. K. *Proc. Fuse/Ammunition/Envir. Sym.* (Battelle Memorial Institute, 1977).
3. Johnson, P. R. *U.S. Army Cold Regions Research and Engineering Lab. Rep.* 77-6 (1977).
4. Napadensky, H. *U.S. Army CRREL Res. Rep.* 119 (1964).
5. Mellor, M. *Snow Mechanics Symp.*, 251 (IAHS Publ. 114, 1975).
6. Chapman, C. R. & Davis, D. R. *Science* **190**, 553 (1975).
7. Sekanina, Z. in *Physical Studies of Minor Planets* (ed. Gehrels, T.) 423 (NASA SP-267, 1971).
8. Housen, K. R., Wilkening, L. L., Chapman, C. R. & Greenberg, R. *Icarus* **39**, 317 (1979).
9. Kresák, L. in *Asteroids*, (ed. Gehrels, T.) 289 (University of Arizona Press, Tucson, 1979).
10. Cintala, M. J. Preprint (Nininger Meteroite Award Paper, 1979).
11. Marsden, B. G., Sekanina, Z. & Yeomans, D. K. *Astr. J.* **78**, 211 (1973).
12. Wilkening, L. L. & Clayton, R. N. *Geochim. cosmochim. Acta* **38**, 937 (1974).
13. Anders, E. *Icarus* **24**, 363 (1975).
14. Shoemaker, E. M. & Helin, E. F. in *Asteroids: An Exploration Assessment* (eds Morrison, D. & Wells, W. C.) 161 (NASA Conf. Publ. 2053, 1978).
15. Roemer, E. *Mem. Soc. R. Sci. Liège*, Ser. 5, **12**, 23 (1966).
16. Chapman, C. R. & Gaffey, M. J. in *Asteroids* (ed. Gehrels, T.) 655 (University of Arizona Press, Tucson, 1979).
17. Wetherill, G. W. in *Asteroids: An Exploration Assessment* (eds Morrison, D. & Wells, W. C.) 17 (NASA Conf. Publ. 2053, 1978).
18. Wetherill, G. W. *A. Rev. Earth planet. Sci.* **2**, 303 (1974).
19. Anders, E. in *Asteroids: An Exploration Assessment* (eds Morrison, D. & Wells, W. C.) 57 (NASA Conf. Publ. 2053, 1978).
20. Kazimirchak-Polonskaya, E. I. in *The Study of Comets* (eds Donn, B., Mumms, M., Jackson, W., A'Hearn, M. & Harrington, R.) 490 (NASA SP-393, 1976).
21. Schaal, R. B. & Hörz, F. *Proc. 8th Lunar Planet. Sci. Conf.* 1697 (1977).
22. Mellor, M. *J. Glaciol.* **19**, 15 (1977).
23. Whipple, F. L. in *Comets, Asteroids, Meteorites: Interrelations, Evolution, and Origins* (ed. Delsemme, A. H.) 25 (University of Toledo Press, 1977).
24. Öpik, E. J. *Ir. Astr. J.* **7**, 141 (1966).
25. Sekanina, Z. in *Comets, Asteroids, Meteorites: Interrelations, Evolution, and Origins* (ed. Delsemme, A. H.) 51 (University of Toledo Press, 1977).

Coseismic changes in groundwater temperature of the Usu volcanic region

Hideki Shimamura & Hidefumi Watanabe

Geophysical Institute, Hokkaido University, Sapporo 060, Japan

Observations of groundwater temperature to an accuracy of 0.001 →C or better have been made at about 30 wells in various areas of impending earthquakes in Japan. The main aim of the observations was to monitor the crustal strain and/or the movement of groundwater which relate to earthquake occurrence. One thermometer was installed at a depth of 200 m in an active volcanic area, Usu, where an upward movement of the magma is occurring. We report here that there was a generally increasing temperature trend at this location: 0.3 →C in the half-year of observation. Decreases in the rate of increase preceded intense local earthquake activity. Temperature jumps were accompanied by large local earthquakes with magnitudes up to 4.2, and the polarity of the steps corresponded to the location of hypocentres. This is probably the first micro-temperature measurement made in the proximity of seismic activity.

Usu volcano, 70 km south-west of Sapporo, erupted on 7 August 1977, after lying dormant for ~30 yr. A series of eruptions followed the first one, continuing until October 1978. Many earthquakes have been observed in the area, beginning just before the first eruption and continuing to the present. The maximum magnitude of the earthquakes was 4.3. Although the earthquake activity has been gradually decreasing since the very first stage of Usu's eruption, it remains at a much higher level than that before the eruption. About two to four earthquakes of magnitude 4 are still occurring every month, so that a lot of small earthquakes are being observed by a local seismological network, consisting of 10 observation sites operated by the Usu volcanological observatory of Hokkaido University.

A precision quartz thermometer was placed in an observation well of depth 380 m, 2 km east of Usu volcano in October 1979. The quartz sensor was installed at a depth of 200 m (below this the well was too narrow). Our thermometer is accurate to 0.001 °C, has long-term stability, and is specially designed for geophysical field observations[1]. Our main aim was to monitor movements of groundwater which were probably related to changes in crustal stress and earthquake generation. Such observation is also interesting from the geothermal viewpoint, as an upward intrusion of magma beneath the volcano has been suggested[2]. Seismologically, this observation is a rare opportunity to observe many earthquakes with magnitude 4 at epicentral distances of less than several kilometres.

In general, the results show a nearly monotonic increase in the groundwater temperature (Fig. 1), which contrasts with observations in other regions; the increment being ~0.3 °C since the start of the observation. However the rate of increment has been slightly decreasing, which probably relates to the gradual decrease in volcanic activity such as crustal deformation or energy release. The thermometer has been tested for long-term stability and found to have a drift rate of less than 0.01 °C yr^{-1}, so that the observed increment cannot be due to instrumental error.

About 30 other thermometers have been installed since 1978 in various areas of impending earthquakes in Japan[1]. However, the Usu thermometer is the only one to have been deployed in an active volcanic area. Hence the general increase in recorded temperature is likely to be due to the upward movement of the magma that has been shown by various volcanological studies[2]. Other thermometers did not record such a large increment in the groundwater temperature, probably because there was no region in which a very shallow magma was rising. For the other wells the d.c.-shifts of the groundwater temperature are <0.03 °C; the present case is more than 10 times larger than those.

Fluctuations in the generally increasing rate are shown in Fig. 1. The decreases in the increasing rate precede high levels of earthquake activity. For example, they are obvious before 13 April, 30 April, and 12 May. Sometimes the ascending trend was even reversed just before intense activity. This applies to periods other than those illustrated in Fig. 1. The intense seismic activities are recurrent, but not periodic; hence we conclude that the 'predictions' of seismic activity by the change in the increasing rate of the observed temperature are caused by some physical mechanism.

It is also interesting that 'seismograms of temperature changes' are frequently recorded by the thermometer (Fig. 2). Almost all large earthquakes were recorded with sharp spikes, and most of them resulted in offsets of the temperature, as is seen in Fig. 2. The onsets of sharp spikes coincide with earthquakes within the accuracy of the sampling interval of the digital thermometer, 15 min; the peaks were reached rapidly, within the same sampling interval of the onset. On the other hand, the average time taken for the spikes to decay was a few tens of minutes. Sometimes the sampling interval was 1 h, as in Fig. 2, instead of 15 min. The amplitudes of the spikes are proportional to the local seismic intensity. The thermometer system was designed to be immune to strong accelerations in the near field of earthquakes and because it was developed to be installed in areas of impending earthquakes, there are no movable parts either in the sensing unit or in the recording unit. It was also tested for exposure to vibrations. We postulate therefore that the recorded spikes were due to frictional heat generated between the pressure case of the sensing unit and the surrounding water in the well as a result of the shaking caused by earthquakes.

Hence the spikes have little geophysical meaning except that we can estimate the thermal time constant with which the generated heat is diffused from the sensing element to the outer groundwater. In contrast, there are substantial offsets remaining

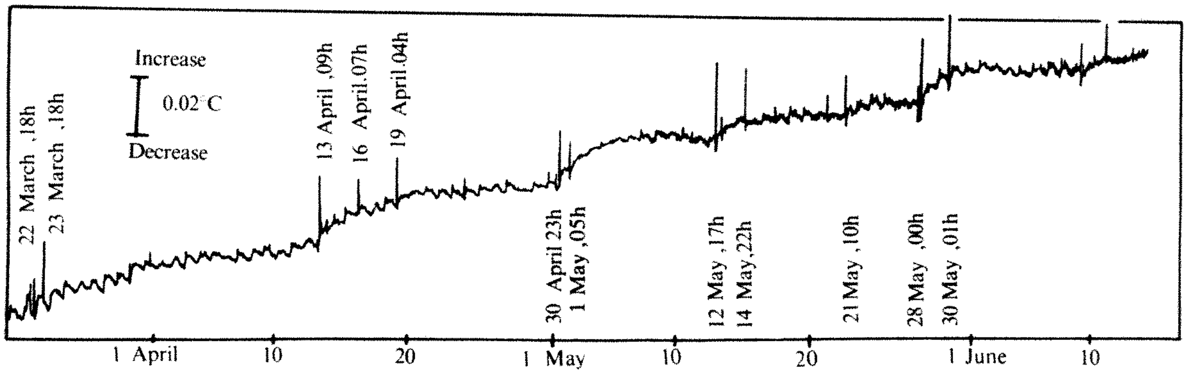


Fig. 1 Record of the groundwater temperature at Usu volcano, obtained by a precision quartz thermometer with an accuracy of 0.001 °C. Spikes correspond to large earthquakes, whose dates and times are indicated.

0028-0836/81/200137—02$01.00

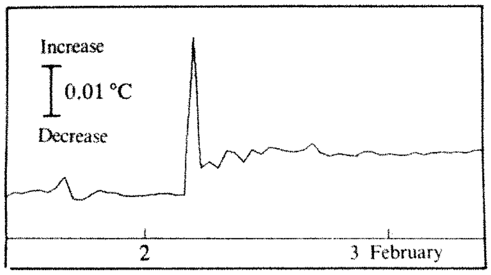


Fig. 2 A 'seismogram' recorded by the quartz thermometer. The record is characterized by a sharp spike which is accompanied by an offset of the groundwater temperature. This record was taken at the sampling interval of 1 h; the leading edges of the spikes actually have steeper onsets. Epicentre, Kita-byobu. Magnitude, 4.1.

Table 1 Magnitude and dates of earthquakes shown in Fig. 3

No.	Date (1980)	Mag	Temperature step (°C)
1	13 April	4.1	+0.0024
2	2 February	4.1	+0.0062
3	8 March	4.2	+0.01
4	23 March	4.1	+
5	23 January	4.2	+0.0063
6	23 March	3.9	−0.002
7	27 February	4.0	−0.0038
8	12 May	4.1	−0.003
9	30 April	3.9	−0.002
10	22 March	4.0	−0.0046
11	24 April	3.9	−0.001
12	14 May	3.9	−0.001

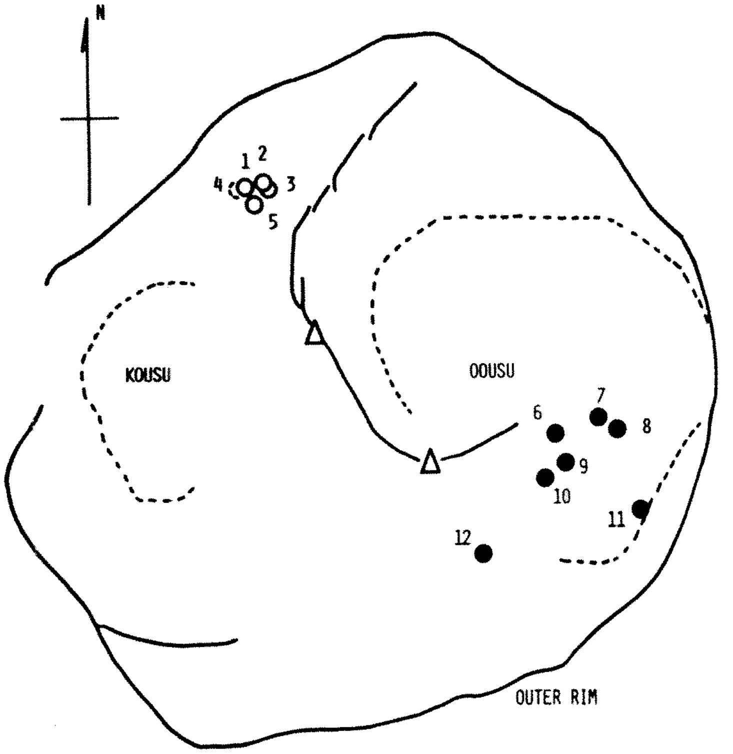


Fig. 3 The observation area, the earthquake epicentres and the location of the observation well. ●, Earthquakes with negative temperature offsets; ○, earthquakes with positive temperature offsets. The U-shaped solid line in the crater is a major surface fault zone. The earthquakes are listed in Table 1. Temperature offsets without a numeral indicate that the amount was too small to be reliably read, although polarity was positive. For magnitudes and dates of earthquakes see Table 1.

after the spikes: both positive and negative offsets were observed which indicates that the recorded offsets are not instrumentally caused by the strong vibration, because the local accelerations of the earthquakes were identical for the two groups which caused inverse polarities. We found that the polarity of the offsets are definitely determined by the locations of earthquake foci; the amounts of the offsets are closely related to earthquake magnitude (Fig. 3). However, we have not found a relationship between the amplitudes of the offsets and the focal depths that were distributed from the surface to 1.2 km.

The largest positive offset, +0.01 °C, was recorded during a magnitude 4.2 earthquake on 8 March 1980. All the group earthquakes (including the magnitude 4.2 earthquakes) that were associated with the positive offsets, occurred beneath the Kitabyobu region, which is the north-west part of the crater of the volcano. The group of earthquakes associated with negative offsets occurred, also without exception, beneath Oousu-minami, which is the south-east part of the crater. The mean epicentral distances of the two groups from the observation well were 2.7 and 1.5 km respectively.

These groups of earthquakes have their own focal mechanisms, which are almost identical within the groups: reverse fault for Kitabyobu, normal fault for Oousu-minami. Both mechanisms have fault planes which are steeply dipping. The fault mechanisms are consistent with a thrust movement, towards the north-east, of the U-shaped surface fault zone, which has been observed by geodetic measurements and geological observations.

At Usu volcano, large earthquakes occur only in these two regions. Smaller earthquakes are more diverse, covering almost all the crater region, and some occur beyond the outer rim, a few kilometres north-east. We did not observe offsets of the groundwater temperature for earthquakes whose magnitudes were <3.7, regardless of their focal locations, thus we do not know whether these small earthquakes generate smaller offsets beyond our detection capability, or whether they inherently do not generate offsets.

We do not have a definite physical explanation for the observed phenomena, the steep trend of the increase in groundwater temperature and the offsets of the temperature. However, the former is probably related to the upward movement of the magma, and the latter to a sudden release of the pore pressure when an earthquake occurs. If so, the monitoring of the pore pressure by the precise monitoring of the groundwater temperature is a promising technique for predicting earthquakes.

Received 24 November 1980; accepted 6 March 1981.

1. Shimamura, H. *J. Phys. Earth.* **28,** 243 (1980).
2. Yokoyama, I., Yamashita, H., Watanabe, H. & Okada, Hm. *J. Volcanol. Geotherm. Res.* (in the press).

Recent glacier variations and volcanic eruptions

Stephen C. Porter

Quaternary Research Center, University of Washington, Seattle, Washington 98195, USA

The injection of volcanic dust and gases into the atmosphere during major eruptions has been advanced to explain short-term variations of climate[1–4]. A calculated global cooling of ~1 K during episodes of intense volcanic activity[3] could lead to a snowline depression sufficient to cause glacier advances equivalent to those of the last several centuries. The hypothesis would therefore be strengthened if a close relationship could be demonstrated between global volcanicity and the pattern of recent glacier variations. Glacier activity in the Northern Hemisphere during much of the past 100 yr shows a lack of synchrony with that in the Southern Hemisphere, but in each hemisphere the sequence of glacier variations matches the acidity record in polar ice cores and the frequency of volcanic eruptions in successive latitude belts. This suggests that glaciers fluctuate in response to atmospheric build-up of volcanic aerosols produced during large eruptions and that explosive volcanism may therefore be a major factor in modulating climate on the decadal scale.

Temperate mountain glaciers are an important source of proxy climatic data, for they respond rapidly to variations of climate by changing in size. A relationship between glacier fluctuations and volcanic episodes would imply a possible link between eruptive activity and climatic change. Fluctuations of many middle-latitude glaciers have been documented during the present century, and for certain glaciers discontinuous observations extend back into the nineteenth, eighteenth or even seventeenth centuries. For glaciers where such data do not exist or are incomplete, records of marginal variations during the past several hundred years often can be reconstructed and dated using geological and botanical methods[5].

Active volcanoes, defined as those which have erupted at least once since AD 1800, are concentrated disproportionately in the Northern Hemisphere; most lie between 60 °N and 10 °S latitude near the margins of converging lithospheric plates[6] (Fig. 1*a*). Few are found south of latitude 20 °S, and their number drops off sharply in both hemispheres polewards of 60°. This asymmetrical global distribution suggests that cumulative long-term loading of the atmosphere by volcanic aerosols may vary with latitude. This inference is supported by Lamb's plot showing the frequency of large dust-producing eruptions as a function of latitude which is dominated by eruptions in equatorial and northern middle latitudes (Fig. 1*b*). Dust entering the upper atmosphere at low latitudes is likely to disperse into both hemispheres, whereas dust injected at high latitudes may remain largely in one hemisphere[1,3]. Consequently, if large volcanic events do affect climate, then differing frequencies of major eruptions in successive latitude belts might produce contrasting climatic responses that could generate asynchronous patterns of glacier behaviour in the two hemispheres.

Most detailed records of glacier variations in the Northern Hemisphere come from mountainous areas lying between 40° and 70 °N latitude (especially Scandinavia, the Alps and northwestern North America) (Fig. 2); although the mountain ranges of central Asia contain many glaciers, sporadic observations, often separated by several decades[7], and the presence of surging glaciers make that region unsuitable for detailed analysis. Glacier advances during the past 100 yr culminated mainly between about 1880 and 1895 and between 1910 and 1925. An intermediate advance, climaxing between about 1900 and 1906, has been reported only south of latitude 62°. An episode of renewed advance began as early as the middle 1950s but increasingly was detected in the 1960s throughout northern middle latitudes. Widespread retreat of glaciers marked the interval between the middle 1920s and the early 1950s, with recession being most pronounced in the 1930s and 1940s[8–11]. Although there were a few reported advances during this interval, most involved surging glaciers. This period of marked retreat, one of the most striking features of the recent Northern Hemisphere glacial record, has attracted much attention, for not only did many glaciers experience considerable reduction in size, but some disappeared entirely.

The few mountain areas at low latitudes that support glaciers are poorly studied, but the limited data available suggest a pattern of fluctuations similar to that of northern middle latitudes.

Southern Hemisphere data come mainly from New Zealand and the Andes, although limited information is also available for certain sub-Antarctic islands. There are fewer recorded observations than for the Northern Hemisphere (Fig. 2) and conclusions based on them are, therefore, less reliable. Despite the large number of glaciers in the Andes, the record from South America is less well known than that from New Zealand and represents a smaller sample. Some of the best-documented advances in the southern Andes involve possible surging glaciers and were excluded from consideration. Glacier advances in the Southern Hemisphere culminated between the middle 1870s and the middle 1890s, between 1906 and 1910, in the first half of the 1930s and in the late 1940s. A recent renewed advance began in the late 1950s to early 1960s, slightly later than in the Northern Hemisphere.

In spite of the common view that glacier fluctuations have been in phase globally during the past 100 yr, in fact they were largely asynchronous between the Northern and Southern Hemispheres for much of this period (Fig. 2). Although a widespread re-advance of glaciers characterized the final decades of the nineteenth century in both hemispheres, a subsequent re-advance in northern middle latitudes that peaked between 1910 and 1925 preceded one in southern middle latitudes which culminated in the early to middle 1930s, a time when Northern Hemisphere glaciers were experiencing their most significant period of retreat. An advance of New Zealand

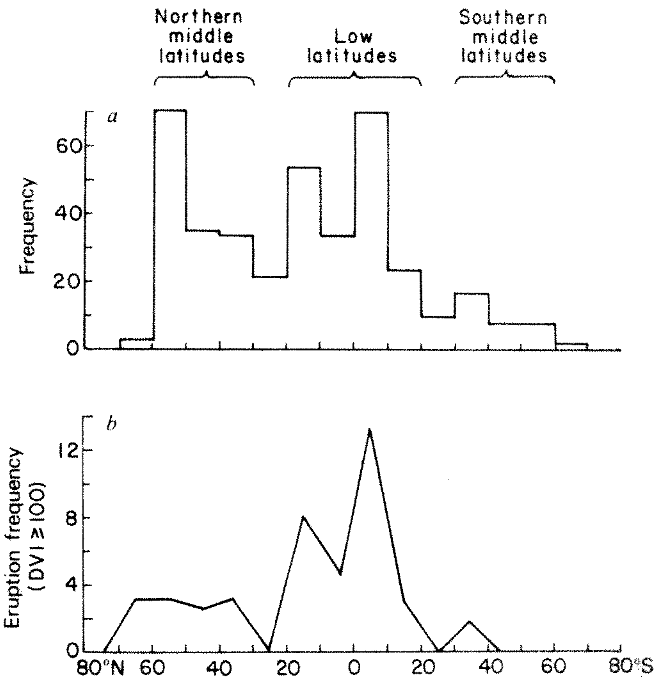


Fig. 1 *a*, Distribution of active volcanoes as a function of latitude. Data compiled from ref. 6. *b*, Number of eruptions per 10° latitude band producing a dust veil index (DVI) of ≥100 (excluding area south of 50 °S) for the period 1800–1964 (from ref. 1, Fig. 19a).

0028-0836/81/200139—04$01.00

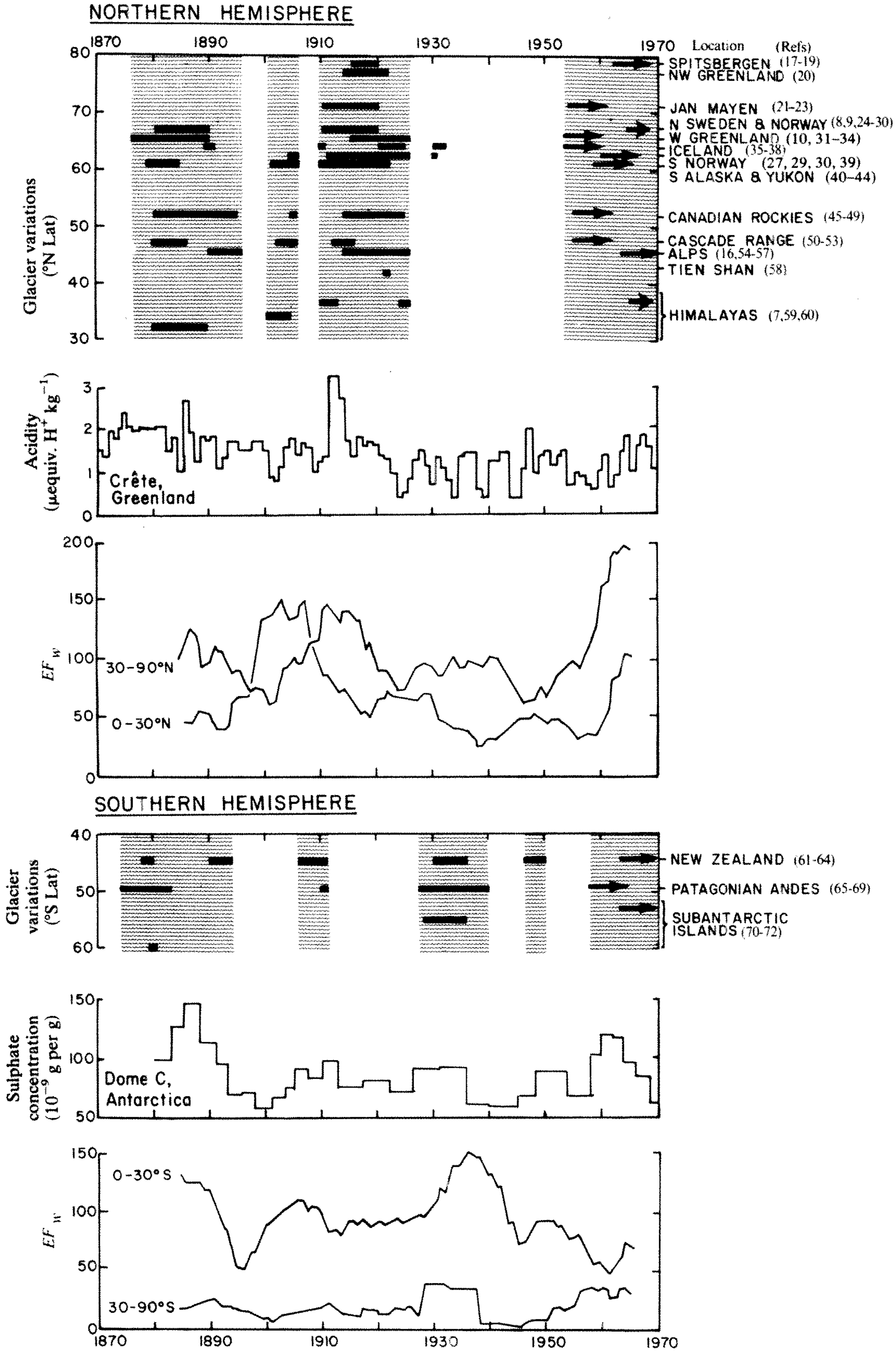


Fig. 2 Comparison of recent glacier variations in the Northern and Southern Hemispheres with acidity profiles from polar ice sheets and frequencies of recorded volcanic eruptions. Data sources for glacier variations given in parentheses. Bars represent intervals within which reported glacier advances culminated and are based mainly on direct observations; in a few cases they are based on ring counts of glacially tilted trees and by lichenometric dating where well-controlled lichen-growth curves have been constructed. Beginning of recent episodes of advance indicated by arrows. Data from known or suspected surging glaciers were excluded because marginal fluctuations of such glaciers may be anomalous and apparently unrelated to local or regional climate. Northern Hemisphere data set includes 115 culminations recorded for 76 glaciers, as well as 10 regional summaries, whereas Southern Hemisphere set represents 50 culminations for 30 glaciers. Spread of dates within each interval mainly reflects response lags of glaciers of different size or having different activity indices, or possible dating uncertainties. Vertically shaded zones show main groupings of data within each hemisphere. Acidity records from Crête (central Greenland) are from ref. 12, while Dome C (Antarctica) data are from ref. 13. Time series of weighted eruption frequencies (EF_w), based on compilation of Hirschboek[14], are 10-yr running totals.

glaciers that culminated in the late 1940s was not matched in northern latitudes where most glaciers continued to recede until the middle 1950s. This contrast in recent glacier fluctuations in the two hemispheres must be explained by any acceptable hypothesis of short-term climatic change.

A comparison of the hemispheric records of glacier fluctuations with variations in the concentration of volcanically-generated sulphate in polar ice sheet cores that has accumulated over the same time interval discloses some striking parallels. Volcanic acids deposited in snow layers shortly after a large eruption can be detected by measuring specific conductivity of meltwater samples or by electrical measurements of solid ice. An acidity profile along the well-dated Crête ice core from central Greenland[12] has strong peaks in the middle 1880s and about 1912–15. Sulphate concentration in snow layers from Dome C in East Antarctica[13] is high in the middle to late 1880s agnitude larger than moderate eruptions. Deca[redacted]ng tals for the period since 1880 are based on more 00 eat and moderate events. The time series show t te 880s and early 1890s were volcanically active in is- heres, but that during the twentieth century the N ern lemisphere experienced periods of high volcanic activity during e 1900s (especially at low latitudes), the 1910s (especially at iddle and high latitudes) and in the late 1950s and the 1960s. olcanoes in the Southern Hemisphere were active during the rst decade of this century at low latitudes, but the major period f eruptive activity in both latitude zones occurred between bout 1928 and 1940.

The two types of volcanic records each have shortcomings for nalysis of the spacial and temporal significance of aerosol-roducing eruptions. The existing ice-core records come from ne polar regions (71.6° N and 74.7 °S, respectively) and prob-bly emphasize eruptive events at middle to high latitudes although certain large low-latitude events like the 1883 erup-ion of Krakatoa (6 °S) are recognizable in both ice sheets; the ower magnitude 1963 Agung event (8.5 °S) is identified as a trong peak in the Antarctic core but is less pronounced in the reenland record). The acidity curves may de-emphasize large ow-latitude eruptions that generate little sulphate but eject fine sh in quantities sufficient to affect the climate for brief periods. The eruption frequency curves do not distinguish between sh-emitting eruptions and important sulphur-rich eruptions, he latter producing the most significant long-lasting atmos-heric effects. Furthermore, the frequency variations must artly reflect varying observational frequency rather than vary-ng volcanic activity. Nevertheless, the similar pattern in the two ypes of record over the past 100 yr suggests that they may rovide a reasonable picture of the temporal and spacial dis-ribution of major eruptions which have had the greatest impact n the atmosphere.

Episodes of glacier resurgence culminated during or shortly after intervals marked by high acidity levels in polar ice and increased frequency of volcanic eruptions in each hemisphere. The high eruption activity of northern middle- to high-latitude volcanoes in the 1910s and a corresponding acidity peak in Greenland ice, for example, coincided with widespread north-ern latitude glacier advances that culminated mainly between about 1912 and 1925, while high volcanic activity, a secondary acidity peak in Antarctic snow, and glacier re-advances marked the late 1920s and the 1930s in southern middle latitudes. The northern low-latitude peak in volcanic activity during the first decade of this century coincides with glacier advances in lower middle latitudes; such advances have not been reported from high latitudes. A similar pattern is seen in the Southern Hemisphere at the end of that decade. The resurgence of glacier activity in both hemispheres beginning in the middle to late 1950s parallels an increase in reported volcanic eruptions, as well as increases in acidity levels in the polar ice sheets.

The data base used in this analysis is of necessity restricted largely to middle latitudes. Moreover, the Southern Hemisphere data are less numerous than those from northern latitudes; this partly reflects the greater extent of land and of glaciated mountains in the Northern Hemisphere, but also a more thoroughly documented observational record. Further testing of the hypothesis, involving information from low-latitude glaciers, must await further data.

Extension of the analysis back in time is also desirable, but the number of carefully-documented glacier observations falls off rapidly in the decades before the middle nineteenth century, and the spacial distribution of data becomes much more restricted. middle latitudes that climaxed between about [illegible]25, and 1845–55[16] match rises in ice-core acidity levels and in the frequency of great ash eruptions[12,14]. These relationships, like those documented above, support the view that volcanic aerosols produced during episodes of large-scale eruptions are a major factor in modulating climate on the decadal scale.

This paper, an outgrowth of research supported by the NSF under grants EAR76-05704 and EAR78-23416, was written while I was a visiting scholar at Scott Polar Research Institute and a visiting fellow in Clare Hall at the University of Cambridge. I thank R. S. J. Sparks and G. de Q. Robin for critical reviews.

Received 10 December 1980; accepted 3 March 1981.

1. Lamb, H. H. *Phil. Trans. R. Soc.* **A266,** 425–533 (1970).
2. Bray, J. R. in *Climatic Change and Variability—A Southern Perspective* (eds Pittock, A. B., Frakes, L. A., Jensen, D., Petersen, J. A. & Zillman, J.) 256–262 (Cambridge University Press, 1978).
3. Pollack J. B. *et al. J. geophys. Res.* **81,** 1071–1083 (1976).
4. Schneider, S. H. & Mass, C. L. *Science* **190,** 741–746 (1975).
5. Porter, S. C. in *Climate and History* (eds Wigley, T. M. L., Ingram, M. J. & Farmer, G.) (Cambridge University Press, in the press).
6. Morris, L. D., Simkin, T. & Meyers, H. *Volcanoes of the World* (World Data Center A for Solid Earth Geophysics, 1979).
7. Mayewski, P. A. & Jeschke, P. A. *Arctic alp. Res.* **11,** 267–287 (1979).
8. Karlén, W. & Denton, G. H. *Boreas* **5,** 25–56 (1976).
9. Theakstone, W. H. *J. Glaciol.* **5,** 411–431 (1965).
10. Weidick, A. *Medd. Grønland* **165,** 1–202 (1968).
11. Eythorsson, J. *Jökul* **13,** 31–33 (1963).
12. Hammer, C. U., Claussen, H. B. & Dansgaard, W. *Nature* **288,** 230–235 (1980).
13. Delmas, R. & Boutron, C. J. *J. geophys. Res.* **85,** 5645–5649 (1980).
14. Hirschboek, K. K. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **29,** 223–241 (1979/1980).
15. Bray, J. R. *Nature* **248,** 42–43 (1974).
16. Le Roy Ladurie, E. *Times of Feast, Times of Famine: a History of Climate since the Year 1000,* 1–426 (Doubleday, New York, 1971).
17. Kosiba, A. *Breslau Univ. Earth Sci. Publ.* B **4,** 1–30 (1960).
18. Kosiba, A. *Bull. Int. Ass. Sci. Hydrol.* **8,** 24–35 (1963).
19. Vivian, R. *Rev. Geogr. Alp.* **52,** 513–531 (1964).
20. Davies, W. E. & Krinsley, D. B. *Int. Ass. Sci. Hydrol. Publ.* **58,** 119–130 (1962).
21. Kinsman, D. J. J. & Sheard, J. W. *J. Glaciol.* **4,** 439–448 (1963).
22. Lamb, H. H., Probert-Jones, J. R. & Sheard, J. W. *J. Glaciol.* **4,** 355–365 (1962).
23. Sheard, J. W. *J. Glaciol.* **5,** 805–811 (1965).
24. Denton, G. H. & Karlén, W. *Quat. Res.* **3,** 155–205 (1973).
25. Karlén, W. *Geogr. Ann.* **55A,** 29–63 (1973).
26. Karlén, W. *Geogr. Ann.* **58A,** 1–34 (1976).
27. Faegri, K. *I.U.G.G. Gen. Assem., Oslo,* 293–303 (1948).
28. Griffey, N. J. *Nor. Geogr. Tidsskr.* **31,** 163–172 (1977).
29. Hoel, A. & Werenskiold, W. *Skr. Nor. Polarinst.* **114,** 1–291 (1962).
30. Karlén, W. *Geogr. Ann.* **61A,** 11–28 (1979).
31. Beschel, R. *Geology of the Arctic* (ed. Raasch, G. O.) 1044–1062 (University of Toronto Press, 1961).
32. Gordon, J. E. *Nature* **284,** 187–189 (1980).
33. Weidick, A. *Medd. Grønland* **158,** 1–196 (1959).
34. Weidick, A. *Bull. Int. Ass. Sci. Hydrol.* **8,** 75–82 (1963).
35. Sigbjarnarson, G. *Jökul* **20,** 50–61 (1970).
36. Thorarinsson, S. *Geogr. Ann.* **25,** 1–54 (1943).
37. Thorarinsson, S. *Jökul* **6,** 1–15 (1956).
38. Thorarinsson, S. *Jökul* **14,** 67–75 (1964).
39. Matthews, J. A. *Norsk. Georgr. Tiddskr.* **28,** 215–235 (1974).
40. Denton, G. H. & Stuiver, M. *Am. J. Sci.* **264,** 577–599 (1966).
41. Field, W. O. Jr. *Geogr. Rev.* **22,** 361–388 (1932).
42. Karlstrom, T. N. V. *U.S. Geol. Surv. Prof. Pap.* **443,** 1–69 (1964).
43. Lawrence, D. B. *Geogr. Rev.* **40,** 191–223 (1950).
44. Viereck, L. A. in *Arctic and Alpine Environments* (eds Wright, H. E. Jr & Osborn, W. H.) 189–204 (Indiana University, 1967).
45. Bray, J. R. & Struik, G. J. *Can. J. Bot.* **41,** 1245–1271 (1963).
46. Heusser, C. J. *Ecol. Monogr.* **26,** 263–302 (1956).
47. Luckman, B. H. *Can. J. Earth Sci.* **14,** 1809–1822 (1977).
48. Luckman, B. H. & Osborn G. D. *Quat. Res.* **11,** 52–77 (1979).
49. West. R. & Maki A. G. Jr. *Science* **133,** 1661 (1961).
50. Burbank, D. W. *Arctic Alp. Res.* (in the press).
51. Harrison, A. E. *J. Glaciol.* **2,** 675–683 (1956).
52. Hubley, R. C. *J. Glaciol.* **2,** 669–674 (1956).
53. Sigafoos, R. L. & Hendricks, E. L. *U.S. Geol. Survey Prof. Pap.* **387B,** 1–24 (1972).
54. Messerli, B., Messerli, P., Pfister, C. & Zumbuhl, H. T. *Arctic alp. Res.* **10,** 247–260 (1978).
55. Patzelt, G. *Z. Gletscherk. Glazialgeol.* **6,** 151–159 (1970).
56. Reynaud, L. *Z. Gletscherk. Glazialgeol.* **13,** 155–166 (1977).
57. Vivian, R. *Les Glaciers des Alpes Occidentales* 1–512 (Allier, Grenoble, 1975).
58. Makarevich, K. G. *I.A.S.H. Publ.* **58,** 249–261 (1962).
59. Mercer, J. H. *Glaciol. Not.* **14,** 19–33 (1963).
60. Batura Glacier Investigation Group *Sci. Sinica* **22,** 958–974 (1979).
61. Burrows, C. J. *N.Z. J. Geol. Geophys.* **16,** 831–855 (1973).
62. Burrows, C. J. & Greenland, D. E. *R. Soc. N.Z. J.* **9,** 321–373 (1979).

Glaciation, erosion and uplift over part of East Antarctica

Peter Wellman & R. J. Tingey

Bureau of Mineral Resources, PO Box 378, Canberra,
Australia 2601

Isostatic rebound effects after the removal of large glacial loads are well known from several parts of the world, but isostatic uplift during glaciation and caused by glacial erosion is less well documented. This is because preglacial landscapes, which provide a datum against which glacial erosion and related uplift is measured, are generally either destroyed by glaciation or not recognized. Remnants of such a preglacial land surface are exposed in the Prince Charles Mountains of East Antarctica. We now report that the amount of glacial erosion here is 1.5 km, measured by the difference in level between the preglacial land surface and the present mean (mostly subglacial) rock surface. The isostatic uplift caused by this erosion is ~0.8 km. Both amounts are thought to be well above average for East Antarctica and glaciated areas in general because the region is traversed by East Antarctica's largest outlet glacier system, parts of which are believed to surge periodically.

The Prince Charles Mountains (72° S, 72° E) are composed of complexly deformed Archaean and Proterozoic igneous, sedimentary and metamorphic rocks which are intruded by Cambrian granites and locally overlain by small intermontane enclaves of Permian sedimentary strata[1]. Near the main glaciers of the Lambert Glacier system the rocks are exposed in large mountains and massifs with broad, roughly concordant, summit plateaus; away from the glaciers there are long low ridges or small nunataks with spiky summits at the same or a lower level[2].

Topographical data[3] and field and airphoto studies have shown that summit altitudes range systematically from 1.25 km in the centre of the area—where summit plateaus are very flat and have 0.2 km of relief—to 2.60 km at the margins where summits, though roughly concordant, are of less uniform altitude (Fig. 1; Fig. 2 curve a). This has led to the interpretation, at first based mainly on morphological grounds, that the summits are relics of a preglacial erosion surface of generally low relief and locally nearly planar. A surface along the flanks of the Amery Ice Shelf with a present altitude of 0.2 km (Fig. 1 stippling) may be a downfaulted continuation of the erosion surface exposed further south. Remnants of nearly planar erosion surfaces of pre-breakup age are widespread in various Gondwanaland fragments[4-6] and we believe that the erosion surface exposed in the Prince Charles Mountains is of the same origin and age.

The summit plateaus in the Prince Charles Mountains, although regarded as the relics of a preglacial erosion surface have, nonetheless, been glaciated and have local mantles of exotic glacial debris and striated pavements. At one locality, however, (Fig. 1D) an erosion surface on banded gneisses is unconformably overlain by a partly vesicular olivine leucitite flows that gives a K–Ar whole-rock age of 50.4 ± 2.0 Myr (ref. 7). The 3 × 1 km flow remnant has vertical sides standing up to 10 m above the adjacent erosion surface. From the existence of this Eocene flow we infer that the underlying erosion is of pre-Eocene and therefore of preglaciation age, and that the

ııo [redacted] kness surveys[illegible] show that the present rock surface
has of up to 4 km. They are used together with the
alti exposed rock masses in estimation of the present
me rface level for 50 × 50 km areas. These in turn were
use alculate a running mean of rock altitudes over 100 × 100 km areas (Fig. 1; Fig. 2, curve e). This mean rock surface (curve e) ranges from −0.5 km along the axis of the lower Lambert Glacier to +0.5 km at the margin of the area studied The difference in level between it and the mean preglaciation surface (Fig. 2, curve a′) ranges from 1.3 to 2.0 km with a mean near 1.5 km; it is a measure of the regional glacial erosion.

Erosion of this order is likely to have caused vertical movement of the crust in the Prince Charles Mountains area although this will have been counteracted by the effect of the present glacial load. The low relief of the preglaciation land surface is interpreted as evidence that the area was stable in relative altitude and hence likely to have been in isostatic equilibrium before glaciation. We concluded[12] that the area is now in

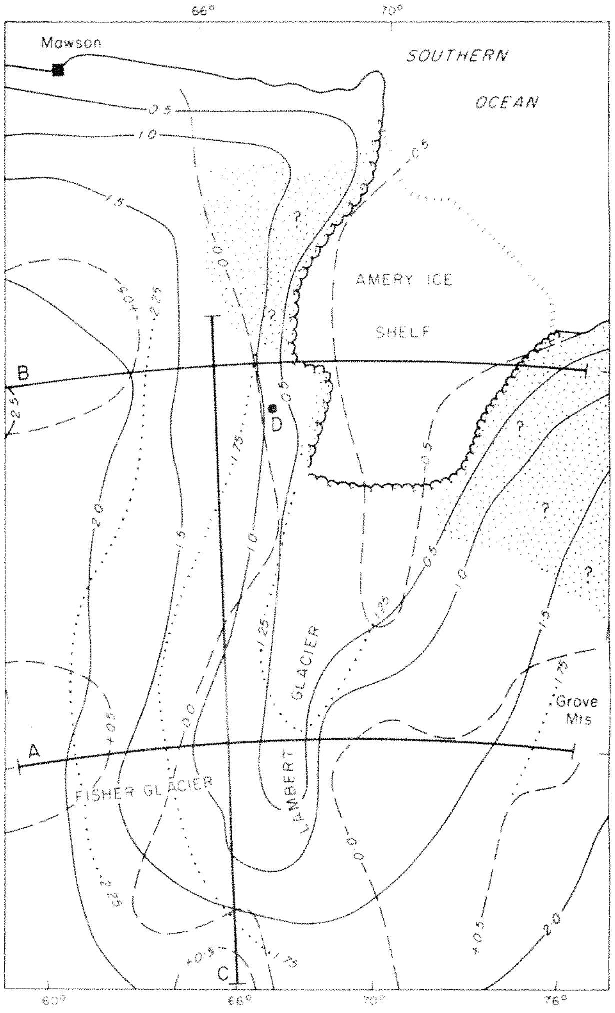


Fig. 1 Map of the Prince Charles Mountains area showing the present-day altitudes in kilometres of the high points of the summit plateau surface (dotted lines), the mean rock surface (dashed lines) and the ice surface (solid lines). Thick lines show the location of Fig. 2 cross-sections. D is the location of the Eocene lava flow discussed in the text. The stippled area is a possible downfaulted portion of the summit plateau surface, now at 0.2-km altitude.

0028–0836/81/200142—03$01.00

isostatic equilibrium from a previous analysis of mean rock altitudes, ice thicknesses for 100 × 100 km areas, and gravity data. If crustal structure does not change, then the removal by erosion of rock masses from a crust in isostatic equilibrium results in an uplift of the crust to maintain this equilibrium. The magnitude of this uplift depends on the density contrast between the crustal rocks and the underlying upper mantle[13]; it is likely to be between 0.80 and 0.87% of the thickness of rock removed. If there were no present-day ice load the original level of the preglaciation peneplain surface could be calculated to have been above the level of the present regional mean rock surface by 0.13–0.20 times the thickness (1.3–2.0 km) of rock removed by glacial erosion, or 0.17–0.40 km.

However, the present mean rock surface (Fig. 2, curve e) is depressed by ice loading, the effect of this loading being the same as that of a rock layer with the same mass per unit area as ice, that is 0.27–0.29 times the ice thickness (Fig. 2, curve b minus curve e) or between 0.3 and 0.5 km. On deglaciation the mean rock surface would be at the level of curve d of Fig. 2. The original level of the preglaciation peneplain would therefore have been 0.17–0.40 km above the level of curve d, at the level of curve c. It would have ranged up to ~1 km altitude near the edge of the region studied. The amount of isostatic uplift that has already occurred as a consequence of glacial erosion is given by the difference between the present level of the mean preglaciation surface (Fig. 2, curve a′) and its original level (curve c). It is between 0.7 and 1.5 km with a mean of 0.8 km.

This uplift and both the rate and amount of glacial erosion that caused it are thought to be above average for East Antarctica and glaciated areas in general because:

(1) The area is traversed by the largest ice stream in the world—there is a large ice flux from a large catchment.

(2) Glacier velocities are high because the area is near the maximum slope of the Antarctic ice cap and the lower sections of the glaciers have basal melting.

(3) At least part of the glacier system may be subject to periodic surging, with consequent short-lived episodes of very vigorous erosion.

However, the relative size of the present-day and past erosion rates is unknown. Modern evidence for near-surface glacial erosion in the Prince Charles Mountains is not abundant and consists of local undercutting of steep cliffs by locally steep, heavily crevassed glaciers. The major ice streams that traverse the Prince Charles Mountains have a base at or near pressure melting point[14], so significant subglacial erosion may be occurring[8,9]. However, it is possible that the valleys are now sufficiently deep that there is no significant erosion at the present flow rate. It seems unlikely that much of the erosion took place during the alpine glaciation that preceded the continental glaciation in the Miocene, because we believe that at this time relief and altitude were low in the Prince Charles Mountains area. It is more likely that erosion along the ice streams has continued during the whole period of continental glaciation but, except initially, has been intermittent and largely related to periodic surging of the glaciers. Such surging is suggested by glaciological studies[14,15], and also from an interpretation of moraines between 0 and 200 m above the present ice level of the Fisher and lower Lambert Glaciers[16]. Normal movement of large glaciers at rates between 50 and 200 m yr^{-1} is achieved by ice shear and recrystallization within the ice mass. During short-lived periodic surges the glacier movement and erosion are several orders of magnitude greater, the fast movement being by the ice moving on a basal water layer.

Evidence for high erosion uplift rates is likely to be found elsewhere in East Antarctica where major glaciers such as the Jutulstraumen traverse mountainous areas near the maximum slope of the ice cap. In mountainous areas of Greenland and north Canada, on the other hand, erosion and erosion-induced uplift are probably occurring more slowly because the glaciers are smaller and zones of high glacier slope are narrower. Nonetheless, the glacial landscapes of the mountains of Greenland and north Canada[8,9] are strikingly similar to those found in the Prince Charles Mountains and are classified as landscapes of selective linear erosion.

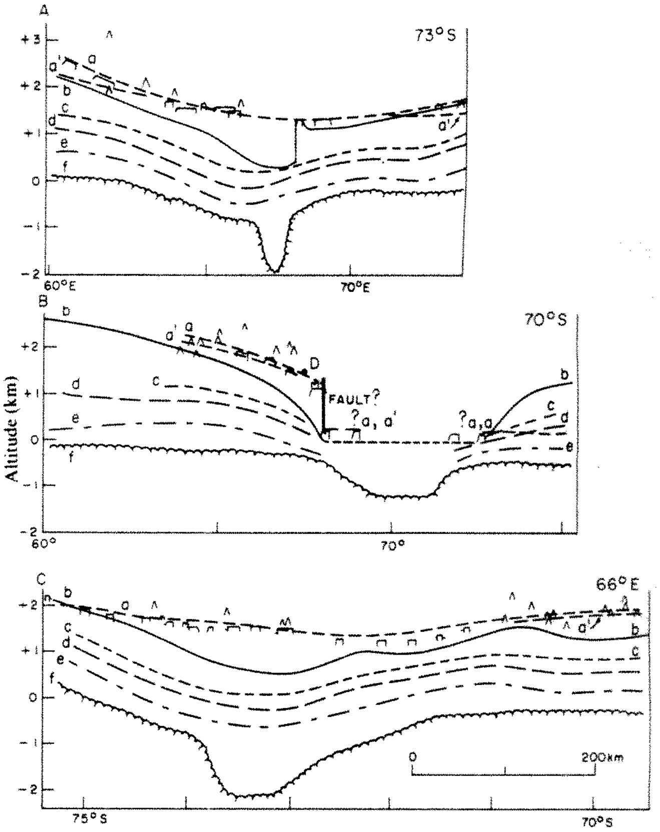


Fig. 2 Sections showing: projected high points of summit plateau and spiky summits, and their mean level (curve a); the present level of the mean preglacial surface (curve a′); the present ice surface (curve b); the original level of the preglaciation surface (curve c); predicted mean level of the present rock surface after deglaciation (curve d); present mean rock surface (curve e); and present lowest rock surface within 100 km of the section line (curve f). Fig. 1 shows the location of these sections.

Erosion rates beneath the main East Antarctic ice cap are probably much lower than those found in the Prince Charles Mountains because of the much lower ice velocity, much less widespread basal melting and lower rock relief. The subglacial landscape in these areas is likely to be one of areal scouring similar to that found in areas of Canada, Greenland and Fennoscandia which are thought to have been once covered by major ice caps[17]. The Fennoscandian and Canadian shields are thought to have been glacially eroded by only a few tens of metres[18].

The average rate of erosion inferred for the Prince Charles Mountains region since the onset of glaciation in the Oligocene[19] is 1.5 km during 30 Myr or 0.05 km Myr^{-1}. This is lower than reported rates[20] for very active orogenic areas: 0.21 km Myr^{-1} for the Himalayas during the past 40 Myr and 0.12 km Myr^{-1} for the Rocky Mountains in the late Cretaceous. It is similar to erosion rates for slightly orogenic areas such as the Appalachians or the Mississippi Basin. Our observations in the Prince Charles Mountains region thus apparently confirm that continental glaciation is not an especially vigorous erosional agent. Furthermore, although the Antarctic continent is widely regarded as being depressed by the ice load there are areas where glacial erosion has induced a net local uplift.

Published with the permission of the Director, Bureau of Mineral Resources, Canberra.

Received 8 December 1980; accepted 27 February 1981.

1. Tingey, R. J. in *Antarctic Geoscience* (ed. Craddock, C.) (University of Wisconsin Press, Madison, in the press).
2. Trail, D. S. *Antarctic Geology* (ed. Adie, R. J.) 143–151 (North-Holland, Amsterdam, 1964).

3. National Mapping, Australia, *1:1,000,000 Base Compilation Maps* SR41–42, SR43–44, SS40–42, SS43–45 (1970); Bureau of Mineral Resources, Australia, *1:250,000 Preliminary Edn Geological Maps* (1972–75).
4. King, L. C. *The Morphology of the Earth* (Oliver & Boyd, Edinburgh, 1962).
5. Truswell, J. F. *An Introduction to the Historical Geology of South Africa* (Purnell, Cape Town, 1970).
6. Idnurm, M. & Senior, B. R. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **24,** 263–277 (1978).
7. Sheraton, J. W. *Bur. Miner. Resour., Aust. Rep.* **194,** 73 (1975).
8. Sugden, D. E. in *Progress in Geomorphology* (eds Waters, R. S. & Brown, E. H.) 177–195 (Institute of British Geographers, London, 1974).
9. Sugden, D. E. *J. Glaciol.* **20,** 367–391 (1978).
10. Morgan, V. I. & Budd, W. F. *J. Glaciol.* **15,** 103–111 (1975).
11. Federov, L. V., Grikurov, G. E., Kurinin, R. G. & Masolov, V. N. in *Antarctic Geoscience* (ed. Craddock, C.) (University of Wisconsin Press, Madison, in the press).
12. Wellman, P. & Tingey, R. J. *Bull. Miner. Resour. J. Aust. Geol. Geophys.* **1,** 105–108 (1976).
13. Woolland, G. P. *Phys. Earth planet. Inter.* **3,** 484–498 (1970).
14. Allison, I. *J. Glaciol.* **22,** 223–235 (1979).
15. Budd, W. F. & McInnes, B. J. in *Climatic Change and Variability: a Southern Perspective* (eds Pittock, A. B., Frakes, L. A., Jenssen, D., Peterson, J. A. & Zillman, J. W.) 228–234 (Cambridge University Press, 1978).
16. Wellman, P. *J. Glaciol.* (in the press).
17. Whillans, I. M. *J. Geol.* **86,** 516–524 (1978).
18. Sugden, D. E. *Geology* **4,** 580–582 (1976).
19. Kemp, E. M. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **24,** 169–208 (1978).
20. Menard, H. W. *J. Geol.* **69,** 154–161. (1961).

Short-term fluctuations in heavy metal concentrations in Antarctic snow

Michael P. Landy* & David A. Peel

British Antarctic Survey, Natural Environment Research Council,
Madingley Road, Cambridge CB3 0ET, UK

Heavy metal variations in polar ice cores may reflect changes in global airborne pollution[1]. Reported profiles from Antarctica[1–3] and Greenland[4], where sampling resolution was normally larger than a year, have shown a much greater variability than can be reasonably accounted for by variations in the annually smoothed emissions of anthropogenic sources and most natural sources. Volcanism has been invoked[4] to account for major features. We now report data from a finely resolved sample sequence from a remote plateau region of the Antarctic Peninsula. Variations of similar magnitude to those reported for longer time series occur also on the scale of seasonal changes and even of single snowfalls; these must be controlled by meteorological processes. We conclude that changes in large-scale transport processes and depositional mechanisms on longer time scales may have as important a role as emission rates in generating the concentration profiles observed in deeper ice cores.

Snow samples were collected in January 1977 from a remote site on the Antarctic Peninsula plateau (70°53′S, 64°57′W; altitude 1,860 m), 170 km from the Weddell Sea to the east and 200 km from the Bellingshausen Sea to the north-west. The plateau at these latitudes is 60–90 km wide and bordered east and west by mountain ranges, though neither was closer than 40 km. The prevailing northerly winds came from an area generally free from exposed rock.

Samples for heavy metal analysis were collected from a clean zone 150 m upwind of a vehicle-free camp site. All sampling tools and containers had been given three lengthy washes in 1 : 6 HNO_3 (Analar) followed by two thorough rinses with ultrapure water in the laboratory before transportation to the sampling site in sealed polythene bags. Two investigators, wearing full clean room garb, dug a 2-m-deep pit with steel shovels and then trimmed 10 cm from the rear wall using a polyethylene scraper.

Triplicated pit-face samples and duplicated 2-m cores (9-cm diameter) recovered directly behind the face with an aluminium hand auger were subdivided with a PTFE saw and stored in 1-l polyethylene containers. Relatively large and erratic metal concentrations subsequently measured in these samples suggested that contamination had occurred, probably from the PTFE saw. However, a 3.4-m core in ~20-cm lengths inside 60-cm acrylic tubes with polyethylene caps gave consistent data with no evidence of contamination; these values are presented here. For the marine ion and microparticle analyses, a 6.5-m core was sealed into polyethylene sleeving and placed inside 100-cm cardboard tubes. The cores were all collected within an area 30-cm square and stored at temperatures below −10 °C before analysis.

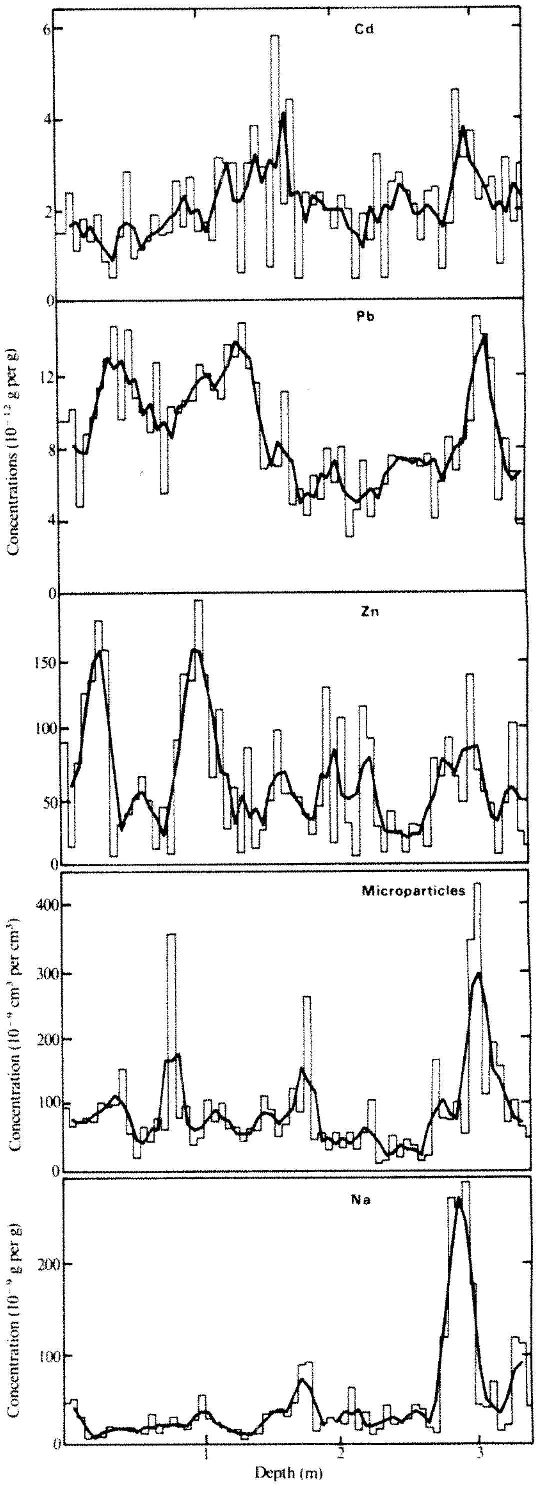


Fig. 1 Concentration profiles for Cd, Pb, Zn, microparticles (0.5–1.8 μm) and Na in the upper 350 cm of snow (May 1973–January 1977). The bold lines are running averages of three data points.

Analyses for Cd, Pb and Zn were performed on snowmelt in a Class-100 laminar flow work station by differential pulse anodic

---

* Present address: Kaiser Aluminum and Chemical Corporation, Center for Technology, PO Box 877, Pleasanton, California 94566, USA.

0028–0836/81/200144—03$01.00

stripping voltammetry (DPASV) at a rotating glassy carbon electrode with *in situ* mercury plating[5]. The segments of the core stored in acrylic tubes were carefully transferred as required into a precleaned polyethylene container kept at −2 °C. The core was briefly removed from its container with insulated PVC clean room gloves. Using a polyethylene scoop, snow was excavated 5 cm into the core, leaving a cylinder of at least 1 cm thickness around the exterior to exclude any superficial contamination. Approximately 30 g of the snow was transferred to a glass voltammetry cell, 50 μl of a mercuric nitrate in 5% HNO_3 plating solution was added (giving 2×10^{-5} M Hg^{2+} at *p*H 2.9) and the sample was melted immediately before analysis. The quantity required for DPASV allowed almost continual analysis every 5-cm down the core, sufficient resolution to permit the study of short-term variations in concentration.

Three DPASV runs were performed on each sample, the final one with a standard addition of 1.5 ng Cd, Pb; 3.0 ng Zn, Cu (15 μl of 10^{-7} g per g Cd, Pb; 2×10^{-7} g per g Zn, Cu in 1% HNO_3). The exact volume of the sample was measured after analysis. Due to the difficulty both of obtaining ultrapure water with heavy metal concentrations consistently smaller than the cleanest snowmelt analysed and of using this to simulate a snow analysis, the concentrations presented here are not corrected for a procedural blank, and as such represent upper limit values. As one core only was available for heavy-metal analysis it was not possible to make a comparison with duplicated samples treated by more complete acid digestion. In a related study at Antarctic sites much closer to exposed rock, we have found[5] that digestion at *p*H 2.9 for 4 h at 110 °C causes no perceptible increase in the observed heavy-metal concentrations except for samples taken very close to and down-wind from rock exposures. Sodium was measured in snowmelt at *p*H 1.5 by flame emission spectrometry, and the microparticle distribution above 0.5 μm was measured in a clean air laboratory using a Coulter counter.

Table 1 Correlation factor matrix, calculated from 3-point running averages

	Cd	Pb	Zn	D*	Na
Cd	1.0				
Pb	−0.03	1.0			
Zn	−0.10	0.29	1.0		
D*	0.19	0.32	0.12	1.0	
Na	0.45	−0.18	0.06	0.32	1.0

*Microparticles in the range 0.5–1.8 μm.

Table 2 Crustal heavy metal enrichment factors calculated from the 3-point running averages in comparison with averages for east Antarctic snow deposited since 1968[1,2]

	Average	Minimum	Maximum	Comparative values for east Antarctic snow
Crustal				
Cd	52.6	14.2	131.5	4,400
Pb	3.5	0.9	7.6	160
Zn	4.0	0.7	12.8	40

The data from the top 3.4 m of snow (66 samples) are summarized in Fig. 1. The overall precision estimated from standards is 20% at the 95% confidence level. Superimposed on the data is a running average of three consecutive values, which reduces any systematic error introduced by comparing data from two different cores where identical snow layers may suffer a relative vertical displacement of up to 10 cm. The annual snow accumulation estimated from both stratigraphic observations and the oxygen isotope ratio profile of a 7-m core taken 100 m from the sampling site was 1.1 m (48 g cm^{-2} yr^{-1}). Support for the stratigraphy is obtained from spectral analysis of the marineion and dust concentration profiles (Fig. 1) which show strong peaks in auto-covariance and cross-covariance with lags of 115 cm. A seasonal periodicity for these species has been reported in other parts of Antarctica[6,7].

The average heavy metal concentrations over the 3-yr sequence are $(Cd = 2.1\pm1.0, Pb = 8.8\pm3.1 \text{ and } Zn = 63\pm47)\times10^{-12}$ g per g. Cu concentrations were usually near or below the detection limit; however, an upper limit of 20×10^{-12} g per g was obtained as an average over 35 samples. These values are similar in magnitude to those published for much more remote sites in Antarctica (Dome C[1] and South Pole[2]), and an order of magnitude smaller than concentrations measured by DPASV on the less remote Adelaide Island (Antarctic Peninsula)[5]. Low concentrations could be the result of a much larger snow accumulation rate at this site compared with the remote sites in East Antarctica, although there is evidence that trace element concentrations in snow are independent of the snow accumulation rate[8].

Previous studies have used more complete digestion of samples before analysis by flameless AAS and therefore these data may contain an additional proportion of metal species not included in our analysis. Our data refer to metal available to DPASV at *p*H 3: free metal ion, labile complexes, metal dissociated at the electrode surface from complexes and colloids, and species adsorbed on mineral matter. From the scant evidence available on the nature and reactivity of heavy metal particulates in the atmosphere, it seems likely that our analysis includes aerosols from the major pollutant sources which include mining, base-metal smelting and refining, and automobile exhausts; these are predominantly oxides, carbonates,

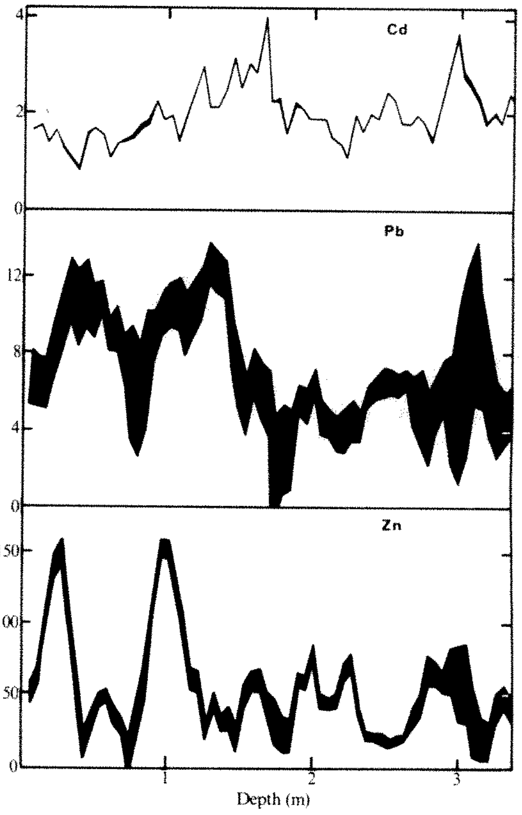


Fig. 2 Microparticle contribution to the heavy metal concentration profiles. The shaded area represents the proportion of metal attributable to microparticles, calculated using average crustal abundances[12].

sulphides, halids and elemental[9]. Emissions of elements and simple compounds by volcanic activity, postulated as a significant source of globally circulating heavy metals[1,3], would also be analysed, along with products of biological processes such as methylation which have been proposed as important natural sources of trace metals in the atmosphere[10,11]. Such emissions would be degraded by photochemical activity during long range atmospheric transport, and probably converted into oxide. Our analysis excludes the contribution from wind-blown dust where heavy metal has replaced other elements in a silicate lattice. Thus we would expect to exclude much of the heavy metal associated with mineral particles derived from aeolian action on the exposed volcanic rocks of the Antarctic Peninsula, which are likely to dominate the mineral content of our samples. The mean dust volume (m^3 per m^3 snowmelt) concentration at our sampling site is 110 p.p.b. (parts per 10^9) which can be compared with a concentration of 24 p.p.b. in the same size range reported[12] for the mean of snowfall at the South Pole deposited immediately before the recent increase in dust concentration at this site due to local human activity[13].

This value may be considered the upper limit to the concentration of microparticles transported to the Antarctic continent from neighbouring continents. On this basis, we estimate that ~80% of the microparticles in our samples are of local origin. If these make a significant contribution to the heavy-metal concentrations observed, then some correlation between the depth profiles for the heavy metals and the microparticle concentrations might be expected. The correlation matrix in Table 1 shows that, with the exception of a weak correlation (significant at the 5% level) between Pb and microparticles which arises from the coincidence of a single major peak at about a depth of 3 m, there was no correlation between the profiles despite the relatively large mineral content.

We can estimate the form of the maximum possible contribution by local dust to the input of heavy metals by assuming that these are of mean crustal composition and that all the heavy metal is available at pH 3. In estimating the mass flux we consider particles in the size range 0.5–1.8 μm. This upper limit prevents a bias in the data due to the influence of a relatively small number of larger particles which are least likely to be available for DPASV analysis. On the average, this size range represents ~53% of the total particle flux between 0.5 and 6 μm diameter. Figure 2 shows how the concentration profiles are modified when allowance is made for the contribution from local dust. The procedure seems to overestimate the contribution because for certain discrete events the estimate of the local heavy metal component is slightly greater than the total amount analysed. Except for occasional periods where trace element concentrations are low and may be almost completely of local origin a maximum average of only 2% of Cd, 29% Pb and 25% of Zn could be attributed to local dust. The residual heavy metal which we interpret as a long-range input to the Antarctic environment, shows enhanced short-period variations.

We have calculated crustal enrichment factors for the heavy metals (Table 2) assuming that the microparticles are of mean crustal composition[14]

$$EF_{\text{crust}} = \frac{\text{[metal] snow/[microparticles] snow}}{\text{proportion metal in average crustal material}}$$

These values represent upper limits as our estimate of microparticle concentration is based on a restricted size range, yet they are between 10 and 80 times lower than enrichment factors calculated for snow in east Antarctica[1,2]. These previous studies have calculated enrichment factors using Al as a crustal reference element, Al being measured[15] in samples of snow melt before digestion for heavy metal analysis. Comparison of the two methods is possible at the South Pole, where Hamilton and O'Kelly[12] measured microparticle size distributions using a Coulter counter in a continuous sequence of samples from a shallow pit covering the period 1956–64. From these data, the authors estimated a mean total dust volume concentration of 24 p.p.b. (72 p.p.b. by weight). This probably underestimates microparticle concentration as particles smaller than 0.5 μm are excluded. If the dust had contained a normal crustal abundance of Al then Al concentrations should have averaged 5.9 p.p.b. For samples covering the same period Boutron *et al.*[2] report direct measurements of 1.5 p.p.b. Al. Therefore, either previous reports represent samples in which Al may have been incompletely analysed due to partial sedimentation of mineral microparticles whilst the samples were melted, or there is considerable error in assuming that the average composition of microparticles penetrating the Antarctic environment is mean crustal in the major elements. This discrepancy could account for a reduction of our EF_{crust} values by a factor of 4, and a further factor of 5 could be due to the previously mentioned dilution effect of local dusts.

Although some of the noise in our concentration profiles may represent variations occurring during individual snowfall events due to progressive scavenging of impurities from the atmosphere, the depth range occupied by some of the major peaks is large compared with the likely precipitation in a single event. Moreover, there is a substantial difference, most noticeable for Zn and Pb, between the mean concentration in the upper 1.5 m and that in the lower 1.5 m on the time scale of seasonal and year-by-year changes. These variations, which range over an order of magnitude for Zn to a minimum factor of 2 for Cd, are much larger than can reasonably be accounted for by variations in global emission to the atmosphere. Sharp variations of similar magnitude, although with a much coarser resolution in longer time scales, have been reported in sample sequences from both the Antarctic and Greenland ice sheets. Several processes, including volcanism[1,3,4,9] have been invoked as possible causes of sporadic variations in global emission rates into the atmosphere.

In support of this idea a 100-yr old sample sequence from central east Antarctica has shown[3] a close parallel between major peaks in the stratospheric burden of volcanic aerosols and the measured concentrations of Zn and Pb. However, a similar mechanism cannot be invoked to account for the shorter term variations observed in our samples which are evidently linked to meteorological transport and depositional processes. As these seem to be capable of generating features in the concentration profiles preserved in snow layers in the short term which are of comparable magnitude to the variations observed in the longer records, we conclude that longer term changes in related meteorological processes may play a significant part in shaping the corresponding record in snow which need not necessarily reflect changes in global emission rates. Until the mechanisms responsible are more fully investigated caution must be exercised in interpreting variations in the concentrations of trace elements in ice cores in terms of varying global emission rates.

We thank J. Mumford for undertaking the analysis of Na and microparticle size distribution in our core samples.

Received 25 October 1980; accepted 13 March 1981.

1. Boutron, C. *Nature* **277,** 551–554 (1979).
2. Boutron, C., Martin, S. & Lorius, C. in *Proc. 9th Conf. Atmos. Aerosols, Condensation and Ice Nuclei*, Galway, Ireland (Pergamon, New York, 1978).
3. Boutron, C. *J. geophys. Res.* **85,** 7426–7432 (1980).
4. Herron, M. M., Langway, C. C., Weiss, H. V. & Cragin, J. H. *Geochim. cosmochim. Acta* **41,** 915–920 (1977).
5. Landy, M. P. *Analyt. chim. Acta* **121,** 39–49 (1980).
6. Thompson, L. G. in *Proc. UGGI Symp. on Isotopes and Impurities in Snow and Ice* (ed. Oeschger, H.) 351–363 (IASH Publ. No. 118, 1977).
7. Herron, M. M. & Langway, C. C. *J. Glaciol.* **24,** 345–356 (1979).
8. Boutron, C. thesis, Univ. Grenoble (1978).
9. Nriagu, J. O. in *The Biogeochemistry of Lead in the Environment*, Part A (ed. Nriagu, J. O.) 155–161 (Elsevier, Amsterdam, 1978).
10. Duce, R. A., Hoffman, G. L. & Zoller, W. H. *Science* **187,** 59–61 (1975).
11. Beauford, W., Barber, J. & Baringer, A. R. *Science* **195,** 571–573 (1975).
12. Hamilton, W. L. & O'Kelly, M. E. *Antarctic Res. Ser.* **16,** 355–362 (1971).
13. Thompson, E. M. & Thompson, L. G. *Antarctic J. U.S.* 136–137 (1977).
14. Taylor, S. R. *Geochim. cosmochim. Acta* **28,** 1273–1285 (1964).
15. Boutron, C. & Martin, S. *Analyt. Chem.* **51,** 140–145 (1979).

Age of mineral equilibria in granulite facies nodules from kimberlites

Ben Harte & Peter M. Jackson

Grant Institute of Geology, University of Edinburgh,
Edinburgh EH9 3JW, UK

R. Mitch Macintyre

Scottish Universities Research and Reactor Centre, East Kilbride,
Glasgow G75 0QU, UK

Among the nodular xenoliths erupted with kimberlites and alkali basalts some plagioclase-bearing and other granulite-facies nodules with metamorphic textures may represent rock samples plucked directly form the Earth's lower crust and Moho region. In studying these xenoliths it is tempting, as in the case of deeper mantle *xenoliths*[1], to use temperature–pressure estimates from mineral compositions to reconstruct petrological–depth relationships and geotherms for the time of eruption.[2–4] In such reconstructions it has been stressed[5–9] that possible modification of the xenolith mineral equilibria during the processes of eruption must be considered; but we emphasize here that the reconstructions may also be invalid because mineral equilibria had become inert or frozen long before eruption. We pay particular attention to nodules from kimberlites in Lesotho, and present some new age data. We argue further that medium- and coarse-grained rocks lacking a fluid phase in stable tectonic regions are unlikely to maintain local equilibrium at ambient geotherm conditions until depths well within the upper mantle are reached.

Granulite-facies xenoliths are relatively abundant in post-Karroo kimberlites occurring close to the supposed position of the southern margin of the Kaapvaal craton in South Africa and Lesotho[4]. Clinopyroxene, garnet and plagioclase are the common major minerals of the xenoliths, with orthopyroxene less common, and amphibole, biotite and scapolite relatively scarce. The nodules often show granuloblastic textures. Rogers and Du Toit[10] noted that the rocks of the nodules were probably related to Proterozoic granulites (of the Namaqua–Natal belt) exposed near the margin of the Kaapvaal craton. Griffin *et al.*[4] partly endorse such a relationship, but they relate the granuloblastic textures to annealing caused by heating during the Karroo episode, and consider the mineral equilibration temperatures to represent deep crustal temperatures at the time of kimberlite eruption. Betton and Cox (see refs 11, 12) suggest that the rocks represented by the nodules were initially formed adjacent to the crust–mantle boundary by crystallization of Karroo magmas and that they reequilibrated in deep crustal conditions before kimberlite eruption.

Recent work[3,4,13–15] has provided a relatively large amount of mineralogical data for estimating temperature–pressure conditions of formation for the granulite-facies nodules from northern Lesotho. For the plagioclase-bearing assemblages: Griffin *et al.*[4] estimate temperatures predominantly between 550 and 770 °C with pressures spreading largely from 4.5 to 15.5 kbar; Jackson[15] estimates temperatues and pressures in the ranges 600–800 °C and 15–20 kbar. These temperatue–pressure estimates yield unexpected results when used to reconstruct the thermal and petrological characteristics of the crust–mantle section underlying Lesotho at the time of kimberlite eruption. The temperatures, except in combination with the highest-pressure estimates, are distinctly higher than those expected by comparisons with normal geotherms for Precambrian shields and platforms and present data for southern Africa[16,17]. Also the higher pressure estimates[15] clearly place plagioclase-bearing rocks within the present day mantle as defined by available seismic data[18] on the depth of the Moho in South Africa.

Assuming that the nodules reflect ambient crustal temperature–pressure conditions at the time (Cretaceous) of kimberlite eruption, Griffin *et al.*[4] attribute the high temperatures estimated to the slow decay of the thermal effects of the earlier Karroo volcanism, and postulate a Cretaceous geotherm with several inflections. We suggest that the possibility must be considered that rocks in the lower crust contain frozen mineral equilibria, which, in this case, may be relics of Proterozoic orogenesis or of high geothermal gradients at the actual time of Karroo magmatism. Although the temperatures of 550–800 °C estimated from the granulite-facies nodules are high for stable continental crust they are by no means high compared with those indicated by amphibolite- and granulite-facies rocks in exposed orogenic belts. Traditionally, rocks collected at the Earth's surface from exhumed orogenic belts are assumed to contain frozen mineral equilibria predominantly reflecting the highest temperatures (or highest entropy states[19]) reached by the rocks. This assumption, which is applied to both volatile-bearing and volatile-free assemblages, is essentially the opposite of that of presumed approximate continual equilibration at depth adopted for the granulite facies nodules in the reconstruction of geotherms for the time of eruption.

The relatively dry nature of the nodule mineral assemblages from Lesotho as shown by the restricted occurrence of primary hydrous minerals (amphibole and biotite)[4,14,15], together with their estimated temperatures of formation, must indicate metamorphism in conditions of highly restricted water content[20]. Where such 'dry' rocks are subjected to metamorphism involving little deformation in the temperature range indicated for the nodules then the products usually fail to approach general textural and mineralogical equilibrium on the scale of a thin section—typically original textures and compositions are partly preserved whilst minerals are clouded and show partial replacement and corona textures[21–24]. Such features contrast with the general appearance of textural and mineralogical equilibrium seen in the Lesothon granulites[4,13–15], and cast doubt on them having equilibrated to ambient geotherm conditions since their initial formation.

We have endeavoured to obtain direct evidence on the age of the mineral equilibria in the nodules by radiometric dating. Two ^{206}Pb–^{238}U ages previously obtained[25] for zircons from the Lesothon granulite xenoliths yielded Proterozoic ages of 1,500 and 1,050 Myr. However, the refractory nature of zircon makes it unlikely that it reflects the time of recent equilibrium in most minerals. Biotite and amphibole are present in some of the nodules and the relatively low blocking or closure temperatures for Ar accumulation imply that the ^{40}K–^{40}Ar method on these minerals provides a means of assessing the minimum age of the mineral compositions yielding the temperature–pressure estimates. We have found only two nodules with sufficient fresh and texturally primary amphibole or biotite for separation and analysis. The K–Ar analytical results for the minerals from these nodules are given in Table 1: the mica yields an apparent age of ~1,000 Myr and the amphibole 714 Myr.

These Proterozoic ages clearly far exceed the < 100 Myr ages measured for emplacement of southern African kimberlites[25]. The two ages may relate to different geological events, but it would have been more reassuring had they been consistent. Furthermore the incorporation of extraneous ^{40}Ar which leads to erroneously old ages is well documented in some metamorphic micas[26] including, in particular, those from granulite facies rocks[27]. However a K–Ar whole-rock age of 90 Myr for the Monastery kimberlite[28] agrees with the Rb–Sr and U–Pb ages[25,29] and therefore does not suggest the presence of any extraneous Ar in the kimberlite. Also the 1,000 Myr biotite age agrees well with the 1,050 Myr obtained for one of the zircons

0028–0836/81/200147—02$01.00

Table 1 K–Ar analyses of minerals from Matsoku Pipe, Lesotho

Sample no.	K (wt %)	^{40}Ar* ($\times 10^{-4}$ cm^{3}g^{-1})	^{40}Ar*/^{40}Ar$_T$	Age† (Myr)
Biotite	7.72	4.063	0.988	1,010 ± 20
(LBM 70)‡		4.013	0.989	1,000 ± 20
Amphibole	1.31	0.4463	0.974	714 ± 14
(LBM 43)§		0.4468	0.978	714 ± 14

* ^{40}Ar = radiogenic argon – 40; ^{40}Ar$_T$ = total argon – 40.

† $\lambda_\beta = 4{,}962 \times 10^{-10}$ yr^{-1}; $\lambda_K = 0.581 \times 10^{-10}$ yr^{-1}; ^{40}K = 1.167×10^{-4} atom %.

‡ Mineral assemblage: garnet–clinopyroxene–biotite–rutile.

§ Mineral assemblage: garnet–orthopyroxene–amphibole–biotite.

extracted from the nodules[25], and is further supported by the extensive occurrence of ~1,000 Myr ages in the Namaqua-Natal Mobile Belt[30-32] on the borders of the Kaapvaal Craton. However, an Rb–Sr measurement on biotite LBM70 (Table 1) yielded ages ($\lambda_{Rb} = 1.42 \times 10^{-11}$ yr^{-1}) of 178 and 200 Myr with assumed initial ^{87}Sr/^{86}Sr ratios of 0.710 and 0.704 respectively.

Amphibole has recently been used quite satisfactorily to date granulite facies metamorphism[33] and extraneous Ar is relatively unknown in this mineral; so, although the 715 Myr age measured on LBM43 is younger than those normally reported for the Namaqua-Natal belt, it could, nevertheless, reflect a Proterozoic event. It is similar to a zircon age of 650 Myr reported from an inclusion in kimberlite from further north-west in Swaziland[25] and to a K–Ar age of 728 Myr measured on primary mica from a peridotite nodule in the Monastery Pipe (R. M. Macintyre and J. B. Dawson, unpublished data). Further data are obviously desirable, but the available radiometric ages clearly suggest that the mineral compositions of Lesothon granulite-facies nodules are more likely to relate to Proterozoic events than the ambient temperature–pressure conditions just before kimberlite eruption.

Considering a section of lithosphere, which is not undergoing significant deformation, magmatic or hydrothermal fluid transport events, the question arises of at what temperature and depth along the geotherm are minerals capable of maintaining chemical equilibrium within reasonably short geological time periods (~10 Myr). More information concerning the time taken for minerals to equilibrate at particular temperatures and other conditions (for example, volatile abundances, amounts of strain) has long been considered desirable. In the present instance, with predominantly anhydrous and unstrained minerals of medium-coarse grain size, equilibration must involve volume diffusion and be relatively sluggish. The general petrologic evidence noted above clearly indicates that the temperature required to produce compositional and textural equilibration in such assemblages must be >700 °C. On the basis of petrological data, Yardley[34] suggests that the diffusion of divalent cations in garnet becomes relatively rapid at ~700 °C; but the common occurrence of Al and Cr inhomogeneities in garnets in garnet–peridotite nodules indicates that general equilibrium requires higher temperatures. Other evidence, associated with exsolution in coarse mineral grains, also indicates the probability of long-term disequilibrium in mantle nodules[35,36]. Using approximate calculations[37] for diffusion distances and rates, the recent experimental data on Ca diffusion in clinopyroxene[38] indicates that temperatures >850 °C are required to equilibrate spherical grains of 5 mm diameter in 10 Myr. Hofmann and Hart[39] note the difficulty of achieving local equilibrium in crystalline peridotite lacking a melt phase at 1,000–1,200 °C; whilst comparisons[8,40] of cation distributions and geothermometers for minerals from peridotite nodules may indicate significant limits to attainment of overall equilibrium by diffusion as temperatures drop to 1,100 °C.

With geotherm temperatures higher than 800 °C and possibly in excess of 1,000 °C necessary for equilibration of essentially volatile-free, magma-free and unstrained mineral assemblages in a tectonically stable lithosphere section, there is a need to reconsider several aspects of the interpretation of upper mantle lithology and their bearing on magma genesis[39]. Using a temperature of 900 °C as a minimum for equilibration by volume diffusion, the geotherms of Clark and Ringwood[41] indicate the persistence of frozen mineral assemblages to lithosphere depths of 60 km beneath ocean basins and 105 km beneath continental shields. It follows that the 'dry' mineral assemblages adjacent to the Moho both in the lower crust and in large tracts of the upper part of the upper mantle cannot be predicted as usually assumed (see ref. 42) by a knowledge of phase equilibria, bulk composition and stable geotherm temperature–pressure estimates. Instead the mineral assemblages will be largely a reflection of past events, presumably dating from when the lithosphere section considered was located at a plate margin or otherwise involved in extensive major deformation and/or heating. The presence of old mantle lithosphere with largely frozen mineral equilibria will add to and emphasize the compositional heterogeneities within the upper mantle and provide an extensive source of material for contamination of magmas. The apparent absence[43] in kimberlites of ulramafic nodules representing the upper part of the upper mantle (at ambient geotherm temperatures <900–1,000 °C) could simply be the result of sampling only frozen mineral equilibria at these depths; with the corollary that ultramafic nodules indicating temperatures of mineral formation around 1,000 °C (±100 °C?) could be derived from the extensive depth range of mantle lithosphere lying at lower geotherm temperatures. Given the existence of such thicknesses of frozen lithosphere the similarities between nodules from Precambrian and Cretaceous African kimberlites[44] become less unexpected.

Received 5 December 1980; accepted 27 February 1981.

1. Boyd, F. R. *Geochim. cosmochim. Acta* **37,** 2533–2546 (1973).
2. Padovani, E. R. & Carter, J. L. in *2nd int. Kimberlite Conf., Extended Abstr.* (1977).
3. Griffin, W. L., Carswell, D. A. & Nixon, P. H. in *2nd int. Kimberlite Conf., Extended Abstr.* (1977).
4. Griffin, W. L., Carswell, D. A. & Nixon, P. H. in *The Mantle Sample: Inclusions in Kimberlites and other Volcanics*, 59–86 (American Geophysical Union, Washington DC, 1979).
5. Boyd, F. R. *Yb. Carnegie Instn. Wash.* **74,** 525–528 (1975).
6. Irving, A. J. *Am. Miner.* **61,** 638–642 (1976).
7. Grew, E. S. in *2nd int. Kimberlite Conf., Extended Abstr.* (1977).
8. Fraser, D. & Lawless, P. J. *Nature* **273,** 220–222 (1978).
9. Mitchell, R. H., Carswell, D. A. & Clarke, D. B. *Contr. Miner. Petrol.* **72,** 205–217 (1980).
10. Rogers, A. W. & Du Toit, A. L. *13th A. Rep. Geol. Comms Cape of Good Hope* (1908) 9–109 (1909).
11. Betton, P. J. & Cox, K. G. in *Geokongres* **79,** Port Elizabeth (1979).
12. Cox, K. G. *J. Petrology* **21,** 629–650 (1980).
13. Rogers, N. W. *Nature* **270,** 681–684 (1977).
14. Jackson, P. M. & Harte, B. in *2nd int. Kimberlite Conf., Extended Abstr.* (1977).
15. Jackson, P. M. thesis, Univ. Edinburgh (1980).
16. Chapman, D. S. & Pollock, H. N. *Earth planet. Sci. Lett.* **28,** 23–32 (1975).
17. Pollock, H. N. & Chapman, D. S. *Tectonophysics* **38,** 279–296 (1977).
18. Hales, A. L. & Sacks, I. S. *Geophys. J. R. astr. Soc.* **2,** 15–33 (1959).
19. England, P. C. & Richardson, S. W. *J. geol. Soc. Lond.* **134,** 201–213 (1977).
20. Stern, C. R., Huang, W. L. & Wyllie, P. J. *Earth planet. Sci. Lett.* **28,** 189–196 (1975).
21. Wilcox, R. E. & Poldervaart, A. *Bull. geol. Soc. Am.* **69,** 1323–1368 (1958).
22. Mason, R. *Norsk geol. Tidsskr.* **47,** 237–248 (1967).
23. Wagner, M. E. & Crawford, M. L. *Am. J. Sci.* **275,** 653–682 (1975).
24. McLelland, J. M. & Whitney, P. R. *Contr. Miner. Petrol.* **73,** 243–251 (1980).
25. Davis, G. L. in *2nd int. Kimberlite Conf. Extended Abstr.* (1977).
26. Brewer, M. S. *Earth planet. Sci. Lett.* **6,** 321–331 (1969).
27. Foland, K. A. *Geochim. cosmochim. Acta* **43,** 793–801 (1979).
28. Macintyre, R. M. & Dawson, J. B. *ECOG IV*, Amsterdam, 66 (1976).
29. Allsopp, H. L. & Barrett, D. R. *Phys. Chem. Earth* **9,** 605–617 (1975).
30. Nicolaysen, L. O. & Burger, A. J. *Sciences Terre* **10,** 497–516 (1965).
31. Kroner, A. *Geol. Soc. S. Afr. Spec. Publ.* **4,** 181–188 (1978).
32. Botha, B. J. V., Grobler, N. J. & Burger, A. J. *Trans. geol. S. Afr.* **82,** 1–5 (1979).
33. Baski, A. K. & Wilson, A. P. *Chem. Geol.* **30,** 109–120 (1980).
34. Yardley, B. W. D. *Am. Miner.* **62,** 793–800 (1977).
35. Harte, B. & Gurney, J. J. *Phys. Chem. Earth* **9,** 367–387 (1975).
36. Dawson, J. B., Smith, J. V. & Hervig, R. L. *Phil. Trans. R. Soc.* **A297,** 323–331 (1980).
37. Kingery, W. D. *Introduction to Ceramics* (Wiley, London, 1960).
38. MaCallister, R. H., Brady, J. B. & Mysen, B. O. *Yb. Carnegie Instn. Wash.* **78,** 574–577 (1979).
39. Hofmann, A. W. & Hart, S. R. *Earth planet. Sci. Lett.* **38,** 44–62 (1978).
40. Fabriès, J. *Contr. Miner. Petrol.* **69,** 329–336 (1979).
41. Clark, S. P. & Ringwood, A. E. *Rev. Geophys.* **2,** 35–88 (1964).
42. Ringwood, A. E. *Composition and Petrology of the Earth's Mantle* (McGraw-Hill, New York, 1975).
43. Dawson, J. B. *J. geol. Soc. Lond.* **134,** 173–184 (1977).
44. Danchin, R. V. & Boyd, F. R. *Yb. Carnegie Instn. Wash.* **75,** 531–538 (1976).

Mutualism in which ants must be present before plants produce food bodies

Stephen J. Risch

Section of Ecology and Systematics, Division of Biological Sciences,
Cornell University, Ithaca, New York 14850, USA

Fred R. Rickson

Botany Department, Oregon State University, Corvallis,
Oregon 97331, USA

Certain interactions between ants and plants can be classified as mutualisms[1–3], with benefit accruing to both members. The plant provides a source of energy, either as solid food or as nectar, and sometimes a domicile such as a hollow stem (or a stem capable of being made hollow by the ants) or hollow stipular thorns. The ants provide the plant with defence against herbivory and/or vine overgrowth. Whether the plant relies entirely or only in part on the ants for its defence, in relationships studied so far it always produces a source of food irrespective of the ants' presence. However, we report here that in the mutualistic relationship between the plant *Piper cenocladum* and the ant *Pheidole bicornis*, the production of food bodies is closely tied to the presence of the ants, so that when they are removed, production of food bodies nearly ceases and when ants reinvade the plant, production begins again. This is the first description of such a relationship in which the food tissue of the plant is induced by the insect.

Risch *et al.*[2,4] have shown that three Costa Rican species of *Piper* share a mutualistic relationship with *Pheidole bicornis* ants. These *Piper* species are slender, shrub-like plants that commonly grow to a height of several metres. They occur in the understory of mature, wet evergreen forests but are most common in more exposed areas such as light gaps caused by fallen trees or along creeks or paths. The plants provide a domicile within a cavity formed by the appressed, curled margins of the petiole (Fig. 1*a*), and the ants obtain additional living space by hollowing out the stem. Apparently the growth of the plants is not affected and no important structural damage is done. In return for the domicile and food tissue, the ants seem to help keep the plant free of invading vines, and remove small herbivorous insects and their eggs.

On the recurved adaxial surface of the petiole margins of *P. cenocladum*, single-celled food bodies are produced. They are collected and consumed by the ants. When a food body forms, the nucleus of an epidermal cell enlarges to seven to eight times its normal diameter and the external cell wall extends outwards (Fig. 1*b*, *c*). The nucleus then moves into the newly formed region of the cell and continues to enlarge together with the rest of the cell. The mature food cell (300–500 μm in diameter; Fig. 2) has a dense cytoplasm rich in protein and is filled with lipid droplets. (Chemical analysis by dry weight shows that food cells contain 10.7% protein, 2.1% carbohydrate and 22.2% lipid.)

Our initial observation was made on living plants of *P. cenocladum*, possessing a weak colony of *P. bicornis* ants, which had been established in a greenhouse at Oregon State University. When, after a short period, the colony died, newly produced leaves developed only a few food cells. This condition persisted in all leaves subsequently produced, although the plants thrived in greenhouse conditions. Because there was no precedent for the necessity of ants to be present for food cells to develop, we investigated the relationship between the insect and the plant in experiments conducted in Costa Rica.

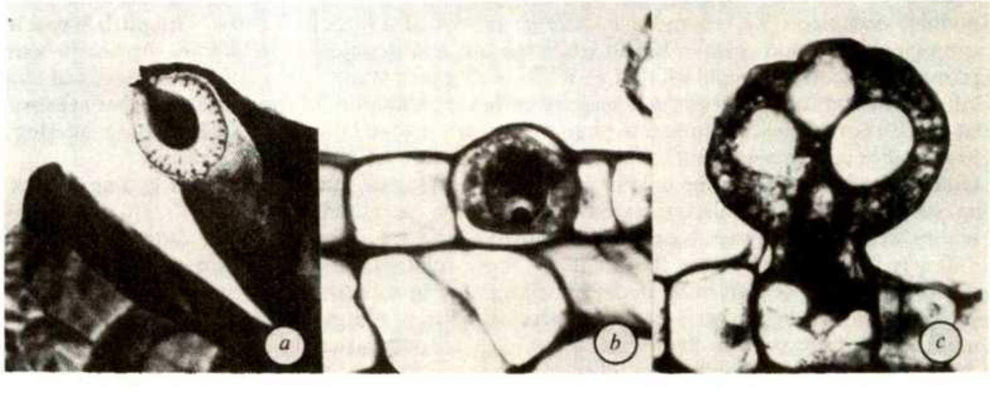
First, we counted the average number of food cells in a *P. cenocladum* petiole tube by dissecting nine petioles from seven plants under a dissecting microscope. The average was 1,500 food cells for a 'typical' petiole. This became the standard against which to compare the results of the ant exclusion experiments described below.

In our first experiment, a 12% aqueous solution of benzene hexachloride (HCH, an insecticide) was injected into the petiole tube of the youngest four leaves of each of nine *P. cenocladum* plants growing at the La Selva Field Station, in the Atlantic lowlands of Costa Rica. After injection, the hole through which ants normally pass from the petioles to the stems was plugged with a sticky resin (Tanglefoot) and the base of the petioles was entirely covered with Tanglefoot to keep ants or other arthropods out of the petioles. After 47 days, all plants were in excellent condition, showing no apparent effects of the HCH. Each treated petiole contained less than 25 healthy food cells, significantly fewer than the number in plants occupied by ants (Wilcoxon signed rank test, $P < 0.01$). There were several large and almost totally decayed food bodies in each of these petioles but virtually no new food bodies were being produced.

In three of the nine plants, active ant colonies were found in petioles below the poisoned area. In those petioles with active ant colonies, the inside of the petiolar sheath was covered with a dense growth of food bodies, similar to that seen in plants not injected. In the remaining six plants, there were no living ants. Although only the top four petioles had received HCH, some of it had probably leaked into the hollow stem and killed the rest of the ant colonies. In these plants, even the petioles below the top four poisoned petioles contained no healthy food bodies.

In our second experiment, ants were added in the field to uncolonized plants of *P. cenocladum* in an isolated population near Turrialba, Costa Rica. By gently prying open the petiolar sheaths, we could see that there were fewer than 50 food bodies in each petiole: we did not actually count them, for that would have destroyed the petioles. *P. bicornis* ant colonies were transplanted to 18 of these plants by attaching an uprooted *P. cenocladum* plant containing an ant colony to each unoccupied plant. (Previous experiments had shown that in these conditions, the colony would usually leave the original plant

Fig. 1 *a*, Basal portion of a *Piper cenocladum* petiole showing the tube formed by the petiole margins. ×2. *b*, Early stage in food cell development. Note the expanding outer wall, enlarged nucleus and dense cytoplasm. ×820. *c*, Mid-stage in food cell development. ×820.



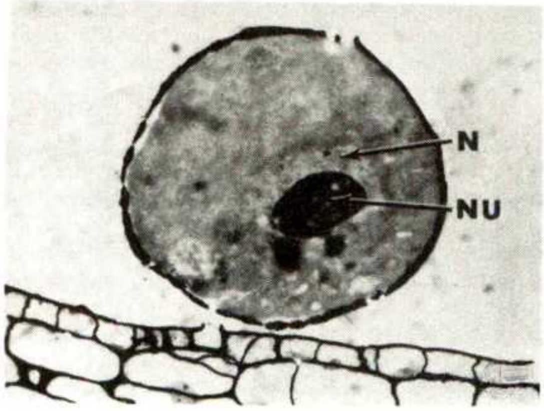


Fig. 2 Mature food body differentiated from a single epidermal cell. N, nucleus; NU, dense nucleolus. ×820.

when it started to wilt and colonize the nearest unoccupied plant.)

After 40 days, 14 of the previously unoccupied plants possessed ant colonies in about one-third of their petioles. Comparison between plants showed that occupied petioles possessed significantly more food cells per petiole ($\bar{x} = 1,020$) than unoccupied petioles ($\bar{x} = 120$) (*t*-test, $P < 0.01$). In other unoccupied plants of the original population, the average was 19 food bodies per petiole, significantly less than either occupied or unoccupied petioles (*t*-test, $P < 0.05$).

In our third experiment, occupied *P. cenocladum* plants were transplanted into containers and moved into a greenhouse at Turrialba, Costa Rica. Nine of them wilted and died back, and their ant colonies departed. After a short period, the plants resumed vigorous growth, but produced no food bodies within the petiole tubes. Of those plants which resumed growth, five were given a new *P. bicornis* ant colony, and four were left unoccupied. After 40 days, nondestructive estimation of the petioles showed that the five with ants were producing a full complement of food cells (>1,000 per petiole), whereas the four without ants had almost no food tissues.

In our final experiment, 12 of the potted *P. cenocladum* plants which had retained their ant colony after transplantation were treated with hydrogen cyanide (HCN) gas. This killed the ant colony in 10 plants: 5 were retained in a greenhouse while the other 5 were further isolated in an indoor laboratory. After 53 days, all 10 plants had virtually no food bodies. Four plants were dissected and examined microscopically. Their petioles contained an average of 15 (1–20) food cells (significantly less than 1,500, Wilcoxon signed rank test, $P < 0.01$). No living arthropods, such as mites, were found in any of these dissected plants. Of the six remaining HCN-treated plants, four received transplanted ant colonies and two were set aside as final controls. After 51 days, all remaining plants were dissected. Those possessing an ant colony contained significantly more food cells per petiole ($\bar{x} = 1,811$ per petiole) than the two control plants ($\bar{x} = 22$ per petiole) (*t*-test, $P < 0.01$).

Each of our four experiments had the same end result: petioles occupied by a colony of *P. bicornis* produced a large complement of food cells (~1,500) while the ant-less petioles possessed a food cell population of 5–50. In addition to stimulating food body production, the ants apparently keep their food source free of fungus. No fungal hyphae were ever observed on food bodies in plants with ants. However, in plants without ants (either found naturally in the field or produced by killing off the ant colonies experimentally), the medium- and large-sized food bodies were nearly always covered with fungal hyphae.

It has been assumed that the plants in these relationships derive maximum benefit in immediately being able to support colonization by the protective ants. Therefore, some energy must always be wasted by the plant in the form of food tissue produced before a colony is established. Such food tissue is either eaten by other animals or falls to the ground. In *P. cenocladum*, however, photosynthate is not used for food tissue before ants arrive. Any energy saved could be put into, for example, the establishment of seedlings or growth. Such energy savings might be especially important to *P. cenocladum* as it grows in the low-light understory of wet, evergreen forests. Three other plants that harbour ants, *Cecropia*, *Acacia* and *Macaranga*, which produce food tissues even in the absence of ants, occur in well lit, disturbed habitats where presumably the amount of photosynthate is not limiting. We have occasionally collected *P. cenocladum* plants containing colonies of *Pheidole* other than *P. bicornis*. There has never been any trace of food bodies in their petioles. Yet in several cases in which *P. bicornis* was found in other petioles of the same plant, those petioles always contained many food bodies. Could it be that the species of *Pheidole* normally occupying *P. cenocladum* is uniquely able to stimulate the production of food bodies in its host?

This work was supported by grants from the Soil and Health Foundation and the NSF.

Received 21 October 1980; accepted 25 February 1981.

1. Janzen, D. H. *Univ. Kans. Sci. Bull.* **47,** 315–358 (1967).
2. Risch, S., McClure, M., Vandermeer, J. & Waltz, S. *Am. Midl. Nat.* **98,** 433 (1977).
3. Rickson, F. R. *Am. J. Bot.* **67,** 285–290 (1980).
4. Risch, S. J. *Brenesia* (in the press).

Voice pitch as an aid to lipreading

Stuart Michael Rosen*, A. J. Fourcin*
& Brian C. J. Moore†

*Department of Phonetics and Linguistics, University College
London, 4 Stephenson Way, London NW1 2HE, UK
†Department of Experimental Psychology, Cambridge University,
Cambridge CB2 3EB, UK

The totally deafened adult, unable to make use of a hearing aid, has no alternative to lipreading for everyday communication. Lipreading, however, is no substitute for hearing speech. Many lipreaders have great difficulty in ideal conditions and even the best lipreaders find the task demanding and tiring. Prosthetic attempts to substitute for lost hearing have centred on three distinct types of intervention, visual[1,2], tactile[3–6] and electrocochlear[7–13]. As none of these is likely to yield a good understanding of speech independent of lipreading in the near future, we have attempted to isolate relatively simple patterns of stimulation that, although not intelligible in themselves, will aid lipreading. From this point of view, the fundamental frequency or 'pitch' of the voice is the most important pattern element because it provides both segmental and suprasegmental information and is practically invisible. It thus complements the visual information already available on the face. As we show here, with the voice pitch presented acoustically, normal listeners can lipread a speaker reading a continuous text at up to two and a half times the rate possible on the basis of lipreading alone. The pitch signal by itself, of course, is completely unintelligible. Although our work is primarily concerned with methods of electrical stimulation of the cochlea[10], it has implications for other sensory substitution techniques, the design of special purpose hearing aids and current theories of speech perception.

Understanding speech solely by vision is difficult for many reasons[14]. On the segmental level (pertaining to individual sounds), some articulatory positions are difficult or impossible to see, for example, g and k. Even for those sounds that are visible, few have unique visual cues. So, 'bet' and 'pet' look alike, as do 'vat' and 'fat'. The most important cause of such confusions is that vocal fold activity, although essential for signalling such contrasts, is invisible. The only difference between 'vat' and 'fat'

0028–0836/81/200150—03$01.00

is that the vocal folds vibrate throughout the v and not the f. In 'bet' and 'pet', the relevant cue is in the timing of the onset of vocal fold activity relative to the opening of the lips. Not only the presence or absence but also the rate of vocal fold vibration, the pitch contour, is important. This information is spread over two or more segments and is thus said to be suprasegmental. Also invisible, the pitch contour can signal among other things, aspects of syntactic structure, whether a question is being asked or a statement made, and which words are stressed in an utterance. Furthermore, voice pitch is the first component of speech mastered by the developing child[15] and is a major indicator of emotional attitude. Although vocal fold activity is not the only information missing in visible facial gestures, it is probably the major one. We should expect provision of it greatly to improve lipreading ability in those who have acquired an auditory knowledge of speech before becoming deaf.

To compare the relative contribution of various types of extra information to lipreading ability, we have made extensive use of a testing technique known as connected discourse tracking (CDT)[3]. In CDT, one person, designated the receiver, must repeat back verbatim what another person, the talker, has said. Every word must be correctly repeated; none may be skipped over. The talker decides on the length of segment to speak at one time although the receiver may interrupt. This length varies both between talker–receiver pairs and within pairs since the length of segments increases as the team become more skilful. When the receiver makes an error, the talker has two basic strategies available. First, and most common, he repeats the bit of text that was received in error. Second, he may comment about the error in any number of ways, for example, give a synonym or paraphrase the meaning. Unlike 'shadowing' tasks, never are the talker and receiver speaking at the same time.

Acoustic (but not visual) isolation between talker and receiver is achieved by placing them in two structurally separate adjoining soundproof rooms with thick viewing windows between and masking noise introduced into the receiver's room. This makes the speech of the talker completely inaudible to the receiver. The talker hears the receiver as normal from a microphone mounted on a set of headphones the receiver wears, routed to an amplifier and loudspeaker in the talker's room. As connected speech, rather than isolated segments, is used, the situation faced by the receiver seems close to that faced by the totally deaf in everyday life. This is important, as some aids, although helpful for isolated segments, show no gain for connected speech[6]. To quantify the task, the talker reads from a novel and performance is measured by calculating the average speed of transmission in words per min over the 5-min testing session.

Using this task we have compared lipreading alone with lipreading plus auditory pitch for five different receivers. Under the lipreading alone (LA) condition, receivers heard nothing through the headphones they were wearing and had to rely solely on visual cues. In the voice pitch (VP) condition, they heard the pitch contour of the talker, after suitable processing in order to make the frequency discrimination capabilities of the normal listeners match those shown by electrically stimulated deaf patients. The pitch of the talker was derived through use of a laryngograph which, by measuring impedance changes, senses vocal fold closure using external electrodes on either side of the speaker's neck[16]. Each receiver had the same talker for all 24 sessions, making a total of 2 hours per condition per talker–receiver pair. Results as a function of session number are shown in Fig. 1, and the mean rates, averaged over all sessions, are given in Table 1. All receivers are considerably faster with the pitch, although the extent of improvement varies. On the whole, the better lipreaders are helped more by the provision of voice fundamental frequency.

For convenience, we have used the laryngograph for the extraction of voice fundamental frequency, but a practical system must work from the acoustic waveform. Preliminary results of ours show that at least one method of acoustic extraction[17], although not nearly as accurate as a laryngographic technique, gives gains in lipreading ability as large as those

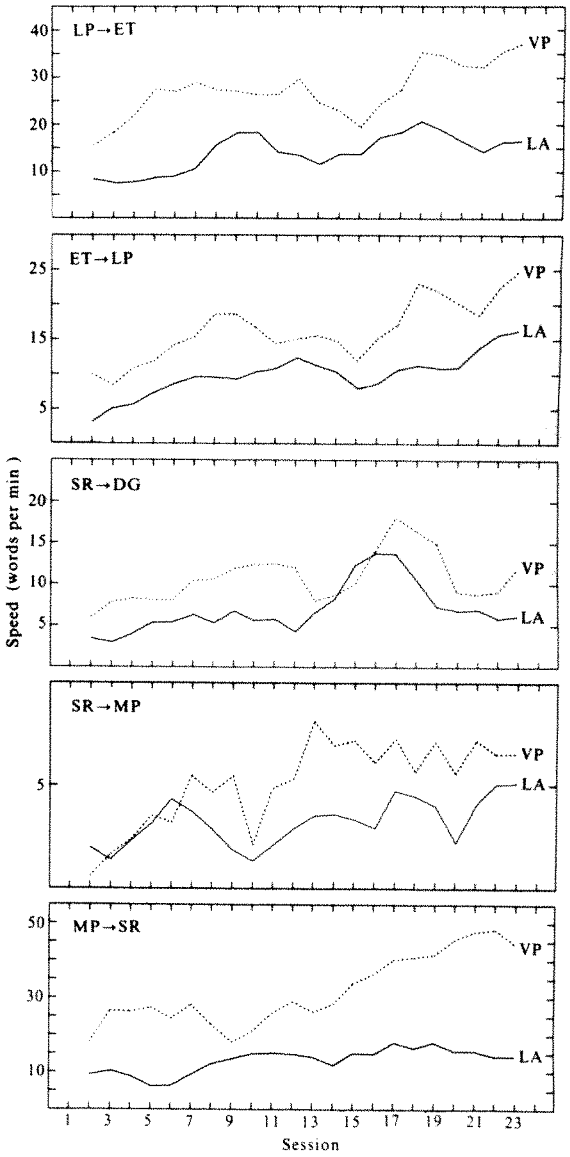


Fig. 1 Speed in words per min as a function of session number for five talker–receiver pairs in a connected discourse tracking task in two conditions. Sessions were 5 min long, making a total of 2 h per condition per talker–receiver pair. These curves result from three-point smoothing of the original data. Note the different scales for the different talker–receiver pairs. In condition LA (lipreading alone), receivers got no auditory information and had to rely totally on lipreading. In condition VP (voice pitch), receivers heard the voice pitch of the talker which was extracted using a laryngograph which senses vocal fold closures by means of a pair of electrodes on either side of the talker's neck. These closures triggered constant-amplitude 20-μs rectangular pulses. Further processing of this pitch information was necessary as electrically stimulated patients show, on the whole, less acute frequency discrimination capabilities than normal listeners. Although the data are sparse and variable on this issue, the most extensive work[21] shows that the median percentage differences in frequency needed for discrimination (the so-called difference limens or DLs) are 4, 7.6 and 6.5% at 125, 250 and 500 Hz, respectively. Patients can achieve values as low as 1–4% over this range[12,21]. We made the normal listening subjects functionally like the deaf by filtering the pulses triggered by the larynograph low-pass at 20 kHz (12 dB per octave) and high-pass at 4 kHz (48 dB per octave) before presenting them to the receiver through headphones (real ear probe frequency response flat within ±5 dB from 300 Hz to 8 kHz) at a peak level of 72 dBA (measured on a flat plate coupler). The receiver still hears the pitch contour of the talker, but in a reduced form. DLs measured for such stimuli on nine young and 'intelligent' but naive subjects were 6.8, 4.4 and 2.1% at 75, 150 and 300 Hz. Our simulation is thus quite good at the low frequencies and still reasonable over the range of frequencies adult males and females (the talkers in the study) typically use (70–350 Hz).

Table 1 Results from a connected discourse tracking task for five talker–receiver pairs averaged over 24 5-min sessions

Talker	Receiver	Speed (words per min) LA	Speed (words per min) VP	% Gain	Absolute gain (words per min)
LP	ET	14.0	27.1	94	13.1
ET	LP	10.0	16.6	66	6.6
SR	DG	6.9	10.7	55	3.8
SR	MP	3.2	5.0	56	1.8
MP	SR	12.9	31.7	146	18.8

Per cent gain is measured by the formula: $100 \times (\mathrm{VP} - \mathrm{LA})/\mathrm{LA}$ where LA and VP are the speeds attained in each condition.

reported above. Other methods of extraction will certainly prove as adequate.

Given that our simulation is a reasonable approximation to the situation experienced by an electrically stimulated deaf patient, these results augur well for single-channel electrocochlear stimulation based on the provision of voice pitch[10]. Comparisons among different methods, electro-auditory and otherwise, will be difficult until all have been tested by similar means. CDT has been used, however, in testing two tactile aids[3,6], giving a percentage gain of the aided over the lipreading alone condition of, at best, 30%. Comparison of such gains with those of Table 1 should give pause to those who maintain that the use of any prosthesis necessitating surgery is premature.

Because pitch, at least when it is presented auditorily, can be of such enormous help in lipreading, it may be fruitful to explore its provision by tactile and visual aids. Initial trials with a single vibrator transmitting pitch information using a temporal code have been disappointing[5], but better results may be obtained when pitch changes are signalled by changes in the place of stimulation[18,19]. Also, there are some patients with residual hearing at usable levels who have such severe recruitment (limited dynamic range) that they are unable to use hearing aids. As intensity changes are not necessary for the use of pitch information, such patients may gain much through use of a hearing aid that extracts and presents the pitch alone at the proper intensity.

Finally, these results contribute more generally to our understanding of speech perception. We have isolated one particular speech feature, presented it to subjects in a totally unnatural way, in a frequency region which, in fact, contains little energy during normally-voiced speech. The subjects have, however, been able to analyse this unfamiliar auditory input with reference to their existing knowledge of speech patterns and combine it effectively with visual information. The possibility that a single pattern feature can be used in this way is a necessary consequence of the point of view that speech perception depends primarily on separable auditory features which are extracted by the listener from the complex acoustic speech stream[20].

This work is done in conjunction with E. E. Douek and G. P. Clarke of Guy's Hospital and J. R. Walliker of University College London and is supported by the MRC, the Sir Halley Stewart Trust and the Royal National Institute for the Deaf. We thank our subjects, David Galbraith, Elisabeth Tunbridge, Lorna Pengelly and Mary Ann Peake; also, the Government Communication Headquarters for the loan of a real-time voice fundamental frequency extractor.

Received 3 November 1980; accepted 18 February 1981.

1. Upton, H. W. *Am. Ann. Deaf* **113,** 222–229 (1968).
2. Martony, J. *et al.* in *Speech Communication Seminar* Vol. 4, 93–96 (Almqvist-Wiksell, Uppsala, 1974).
3. De Filippo, C. L. & Scott, B. L. *J. acoust. Soc. Am.* **63,** 1186–1192 (1978).
4. Kirman, J. H. *Psychol. Bull.* **80,** 54–74 (1973).
5. Rothenberg, M. & Molitor, R. D. *J. acoust. Soc. Am.* **66,** 1029–1038 (1979).
6. Sparks, D. W., Ardell, L. A., Bourgeois, M., Wiedmer, B. & Kuhl, P. K. *J. acoust. Soc. Am.* **65,** 810–815 (1979).
7. Chouard, C. H. & MacLeod, P. *Laryngoscope* **86,** 1743–1751 (1976).
8. Clark, G. M. *et al. J. Lar. Otol.* **91,** 935–946 (1977).
9. Eddington, D. K., Dobelle, W. H., Brackmann, D. E., Mladejovsky, M. G. & Parkin, J. L. *Ann. Otol. Rhinol. Lar.* **87,** Suppl. 53, 1–39 (1978).
10. Fourcin, A. J. *et al. Br. J. Audiol.* **13,** 85–107 (1979).
11. House, W. F. *Ann. Otol. Rhinol. Lar.* **85,** Suppl. 27, 1–93 (1976).
12. Merzenich, M. M., Michelson, R. P., Schindler, R. A., Petit, C. R. & Reid, M. *Ann. Otol. Rhinol. Lar.* **82,** 486–503 (1973).
13. Simmons, F. B. *Archs Otolar.* **84,** 22–76 (1966).
14. Jeffers, J. & Barley M. *Speechreading* (*Lipreading*) (Thomas, Illinois, 1978).
15. Fourcin, A. J. in *The Development of Communication* (eds Waterson, N. & Snow, C.) 47–72 (Wiley, Chichester, 1978).
16. Fourcin, A. J. & Abberton, E. *Med. biol. Illust.* **21,** 172–182 (1981).
17. Gold, A. A. & Rabiner, L. R. *J. acoust. Soc. Am.* **46,** 442–448 (1969).
18. Spens, K.-E. *Q. Prog. Status Rep.*, 2–3 (Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, 1975).
19. Stratton, W. D. *Volta Rev.* **76,** 26–35 (1974).
20. Fourcin, A. J. in *Language, Cognitive Deficits and Retardation* (ed. O'Connor, N.) 33–46 (Butterworth, London, 1974).
21. Bilger, R. C. *Ann. Otol. Rhinol. Lar.* **86,** Suppl. 38, 1–176 (1977).

Magnetic sense of direction in woodmice for route-based navigation

Janice G. Mather & R. Robin Baker

Department of Zoology, University of Manchester,
Manchester M13 9PL, UK

Orientation to the Earth's magnetic field has been shown for bacteria, planarians, molluscs, insects, elasmobranch fish, salamanders and birds[1,2]. Recent work indicates that humans may also have a magnetic sense of direction[3,4]. We present here the first evidence for such a sense in a mammal other than man, the European woodmouse. In addition, by manipulation of the magnetic information available during the outward journey (an approach used previously with pigeons[5,6]), we present the first clear evidence that both route-based and location-based mechanisms can be involved in goal orientation following displacement.

All experiments were conducted on the European woodmouse, *Apodemus sylvaticus*, between 10.00 and 18.00 GMT from 23 June to 23 July 1980, at Woodchester Park Field Centre, Gloucestershire, UK. After capture in a Longworth trap, each individually fur-clipped mouse was allowed to enter an adjustable translucent plastic scroll which was then placed inside a transportation cage (Fig. 1). This cage was air tight and light proof, and supported two lateral coils of copper wire which could be activated to manipulate the ambient magnetic field. After a 2-min adaptation period, the cage plus animal was carried in a straight line roughly 40 m due north to a test site, the animals spending approximately 4 min inside the transportation cage. At this site the induced field was deactivated before the transportation cage was uncovered, taking care to ensure that the position of the Sun could not be detected by the mouse during the period from activation to deactivation of the artificial magnetic field. At the test site, the tube containing the animal was removed from the transportation cage and the mouse allowed to enter an orientation cage, which was housed inside a large cardboard box (70 cm diameter and 60 cm high). This box restricted both visual and olfactory cues, the latter being reduced further by the addition of Perspex doors to the orientation cage. The cage was positioned so that its four arms were directed towards magnetic north, south, east, and west. Four guy ropes anchored the cardboard screen firmly to the ground. The method was as described by Mather and Baker[7]. The cardboard box lid was folded over to create a 25-cm diameter window above the orientation cage. Consequently, only the sky directly overhead could be seen by the mouse from within the cage, any features on the horizon being screened by the box. During roughly half of the tests the sun shone into the box, though the Sun's disk could rarely be seen directly by the mouse in the cage.

Testing consisted of recording the time spent in each of the four arms of the orientation cage during a 4-min test period, observation being through small peep holes in the cardboard

0028–0836/81/200152—04$01.00

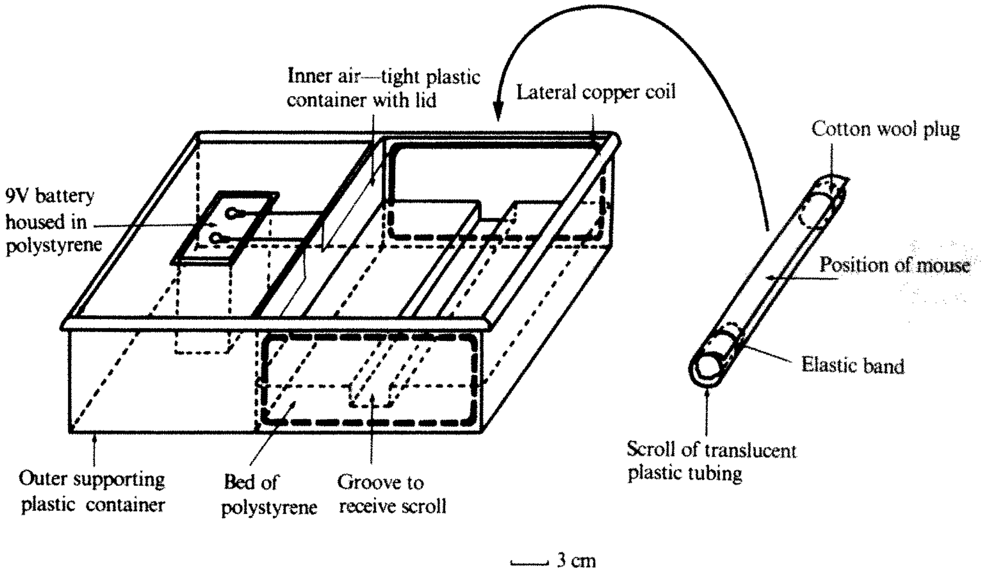


Fig. 1 The transportation cage consisted of a translucent air-tight plastic container (18 × 18 × 9 cm), the inside of which supported lateral coils made from 100 turns of 20 standard wire gauge insulated copper wire, and powered by a 9 V battery. An 82-Ω resistor in the circuit resulted in a current of 65 mA, producing a magnetic field of ~0.37 G between the coils. When this field was directly opposed to the horizontal component of the Earth's normal magnetic field, a near mirror image of the local geomagnetic field was produced. The vertical field at the animal's position was changed by less than 5% whereas the horizontal field was reversed and brought to within 115% of the natural value. This gave a total field strength of 0.52 G. As a result, the magnetic field through the experimental animal could be adjusted so that when the cage was orientated along the north–south axis (as it was during experimental displacement of the test animals), the horizontal component of the ambient magnetic field could be shifted by 180°. Alternatively, the battery could be disconnected so that an animal in the centre of the deactivated transportation cage would experience the Earth's unchanged magnetic field, and could therefore serve as a control. The animal was retained in a plastic scroll, the diameter of which could be adjusted by loosening two supporting elastic bands. This ensured both that the mouse remained in the same orientation (facing direction of travel) during transport to the test site, and that the tube was loose enough for the animal to suffer no ill effects. A cottonwool plug at either end retained the animal in the central part of the scroll, which was then placed in a groove in a bed of polystyrene inside the transportation cage.

box. (We have shown previously that observer position has no significant effect on the orientation response of rodents in this design of experiment[7]. In fact, the present investigation had a built-in control in that as long as observer position remained constant following both control and experimental displacements, any influence of observer position could not account for differences in the orientation of the two groups. Nevertheless, the observers' location was always neutral (at 90°) to the home direction and changed from east to west half-way through the experimental series.) After each test, the orientation cage was covered by a black polythene bag and used to return the animal to its point of capture.

Twenty different animals were tested (13 males), no animal being used on two consecutive days[7]. The first series of experiments consisted of 33 separate displacement occasions, 16 of which used a reversed magnetic field (Fig. 1) during the outward journey. Another 17 displacements served as controls, the coils being disconnected during displacement, the animals thereby experiencing the Earth's normal magnetic field. The influence of smell of previous occupant was controlled by maintaining the orientation cage in the same physical orientation (the same arm pointing north) during testing, and arranging the tests in a sequence of two 'activated' displacements followed by two controls.

Statistical analysis was carried out in two stages, as described by Batschelet[8-10]. Time spent in each of the four arms of the orientation cage was used in a first-order analysis to calculate the mean vector (r, θ) for each 4-min test, where r is the length and θ the mean angle. The mean angles were then subjected to second-order analysis[10]. Non-uniformity was examined by Rayleigh's Z-test[8], and clustering around a predicted direction by the V-test[9]. The Watson $U^2_{n,m}$ non-parametric two-sample test (test statistic U^2)[8] was also used. Statistics are presented with home direction as 0°.

Table 1 Orientation performance in the test apparatus; summary of second-order analysis

Variable	No. of tests	Mean angle (θ)	Mean vector (r)	Z	u (versus home 0°)
Normal field plus screen	17	+14°	0.416	2.940	2.353*
Reversed field plus screen	16	+131°	0.505	4.088†	−1.876
Reversed field no screen	20	+7°	0.483	4.662*	3.031‡

Z = Rayleigh's test, u = V-test.
*$P < 0.01$; †$P < 0.05$; ‡$P < 0.005$.

The results (see Table 1 and Fig. 2) show that there is a marked dichotomy in the directional preferences of the control and experimental animals, the distributions of direction for the two groups being significantly different ($U^2_{17,16} = 0.402$; $P < 0.001$). After transportation involving minimal exposure to visual and olfactory cues, but in the presence of the Earth's normal magnetic field, the mice showed homeward orientation (V-test, Table 1) despite the continuing restriction of these cues by the screen and perspex doors at the test site. When the magnetic vector was deflected by reversing the horizontal component of the magnetic field during displacement, subsequent orientation by the mice at the test site was shifted away from home by 131°, resulting in a 117° difference in the mean orientation of the two groups. Experimentals are significantly clustered about 180° from home ($u = 1.876$, $P < 0.05$), but are not quite significantly clustered about 180° from the mean angle for controls ($u = 1.298$, $P < 0.10$).

It has been suggested that the ability of animals to detect a magnetic field may be associated with deposits of magnetite which have been found in a variety of organisms[2]. In vertebrates, magnetic remanence has been found most often in the head region. We have undertaken pilot investigations into the possibility that similar deposits may exist in the rodent head.

Three species of rodents were used, the European woodmouse, *A. sylvaticus*, the bank vole, *Clethrionomys glareolus*, and the house mouse, *Mus musculus*, giving a total of 12 animals which had been deep frozen for up to 3 months before use. Each specimen was then embedded in a small block of deionized ice and retained in a plastic container for insertion into a par–astatic magnetometer. The heads were tested whole for natural remanence (NRM, a net field due to a preferential alignment among individual magnets), and inducible remanence (IRM, a measure of the total amount of magnetic material). Immediately before testing for IRM, the specimens were exposed to a

2,000-G field to align any permanently magnetic material. The heads were subsequently dissected, glass instruments being used to prevent possible contamination from metallic tools, and the various parts were retested for IRM (see Table 2).

All samples exhibited some NRM and even greater IRM, which was strongest in the anterior part of the head in an area associated roughly with the olfactory region but not directly with brain tissue. The level of magnetic remanence found in rodent heads is comparable with that found in pigeons (10^{-6}–10^{-5} e.m.u.)[11] and bees (10^{-6} e.m.u.)[12], but less than that found in the dolphin (10^{-5} e.m.u.)[13].

The fact that mice, which had experienced a reversed magnetic field during displacement and been allowed no visual or olfactory cues en route and only limited olfactory cues at the destination, orientated away from home suggests that route-based navigation (estimating home direction by monitoring the outward journey)[4] was taking place even in the absence of visual cues, and the involvement of a magnetic sense of direction in this ability. Alternatively, there could be short-term after-effects of exposure to altered magnetic fields. If so, fields experienced during transportation might affect subsequent navigation at the test site. However, we consider this interpretation improbable, particularly as the natural and artificial field strengths were so similar that it seems unlikely for the two to have differing effects.

Having found that we could disrupt the initial orientation response, by restoring visual and olfactory cues at the destination, a study of the involvement of these cues in location-based navigation (perceiving cues at the release site and from these determining home direction)[4] was possible. After the first series of experiments had been conducted, a second series was carried out in which the magnetic field was reversed during displacement, as before, but no screen (cardboard box) was placed around the orientation cage at the test site. All other procedures were identical except that the observers were sited 5 m away. In all, 20 different individuals (15 males) were tested once each (5 had been used in the first series of experiments), over a period of 6 days.

The results are presented in Fig. 2*c* and show that when animals experience a reversed magnetic field during displacement but are allowed relatively unrestricted access to visual cues at the test site, they orient not far away from home, as in Fig. 2*b*, but towards home (Fig. 2*c*). Comparison of the first-order mean angles of the two groups, which both experienced a reversed magnetic field during the outward journey, shows a significant difference ($U^2_{20,16} = 0.273$, $P < 0.01$), whereas the preferred directions of those animals experiencing a normal field plus screen and a reversed field but no screen are not significantly different ($U^2_{20,17} = 0.063$, $P > 0.1$).

Table 2 Summary of natural remanence (NRM) and inducible remanence (IRM) in various regions of the rodent head

		n	Mean magnetic moments ($\times 10^{-6}$e.m.u. g^{-1})	Range
Container				
Empty	NRM	4	0.63	0.28–1.05
Plus ice	NRM	4	0.83	0.59–1.23
	IRM	4	0.96	0.51–1.45
Whole head	NRM	11	1.33	0.29–7.81
	IRM	12	6.31	1.14–18.31
Head (IRM)				
Cranium		3	3.08	1.20–5.57
Lower jaw		7	0.75	0.43–1.64
Cranium (IRM)				
Anterior		6	5.13	1.26–16.13
Posterior		9	1.23	0.59–2.55
Anterior cranium (IRM)				
Anterior		5	1.93	0.47–4.22
Posterior		5	0.82	0.27–2.01
Dorsal		6	1.92	0.71–4.01
Ventral		6	1.31	0.71–2.99

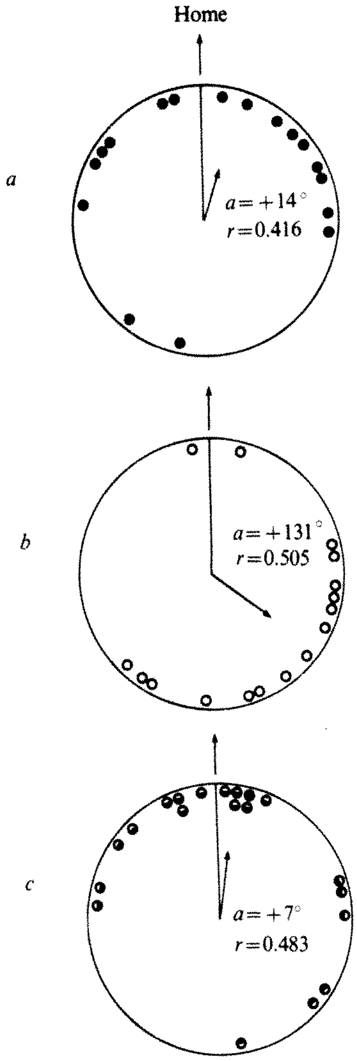


Fig. 2 Effect of conditions during displacement and at the release site on the goal orientation of woodmice. Each dot represents the mean angle (θ, first-order analysis) for an individual during a 4 min test period. Arrows within the circles show the second-order mean vectors of the estimates. Statistics show the mean angle ($a°$) and the length (r) of the mean vector. Angles are expressed with home direction 0°. *a*, Normal field plus screen; *b*, reversed field plus screen; *c*, reversed field, no screen.

Our experiments have shown that: (1) woodmice do have a magnetic sense of direction; (2) they can sample the magnetic field during experimental displacement and can use this information for route-based navigation; and (3) the magnetic information collected during the displacement journey may become subservient to other cues obtained at the test site if the information from the two sources conflict. We stress, therefore, the importance when studying route-based navigation of restricting cues available at the test site, and when investigating location-based navigation, of manipulating in a predictable way the animal's ability to monitor its outward journey.

We thank Dr P. Dagley and Professor F. Oldfield for advice on use of the magnetometer, Dr D. W. Yalden for woodmice specimens, and Mr and Mrs R. Kelly for their hospitality during the field work. This investigation was carried out while J.G.M. was receiving a SRC grant.

Received 29 August 1980; accepted 17 February 1981.

1. Schmidt-Koenig, K. & Keeton, W. T. (eds) *Animal Migration, Navigation and Homing* (Springer, Heidelberg, 1978).
2. Gould, J. L. *Am. Sci.* **68**, 256–267 (1980).
3. Baker, R. R. *Science* **210**, 555–557 (1980).
4. Baker, R. R. *Human Navigation and the Sixth Sense* (Hodder and Stoughton, London, 1981).
5. Wiltschko, R. & Wiltschko, W. *Naturwissenschaften* **65**, 112–113 (1978).

6. Kiepenheuer, J. in *Animal Migration, Navigation and Homing* (eds Schmidt-Koenig, K. & Keeton, W. T.) 135–142 (Springer, Heidelberg, 1978).
7. Mather, J. G. & Baker, R. R. *Nature* **284,** 259–262 (1980).
8. Batschelet, E. *Statistical Methods for the Analysis of Problems in Animal Orientation and Certain Biological Rhythms* (American Institute for Biological Science, Washington DC, 1965).
9. Batschelet, E. in *Animal Orientation and Navigation* (eds Galler, S. R., Schmidt-Koenig, K., Jacobs, G. J. & Belleville, R. E.) 61–91 (GPO, Washington DC, 1972).
10. Batschelet, E. in *Animal Migration, Navigation and Homing* (eds Schmidt-Koenig, K. & Keeton, W. T.) 3–24 (Springer, Heidelberg, 1978).
11. Walcott, C., Gould, J. L. & Kirschvink, J. L. *Science* **205,** 1027–1029 (1979).
12. Gould, J. L., Kirschvink, J. L. & Deffeyes, K. S. *Science* **201,** 1026–1028 (1978).
13. Zoeger, J., Dunn, J. & Fuller, M. *EOS* **61,** 225 (1980).

Stimulation of aldosterone production by β-melanotropin

Hiroaki Matsuoka*‡, Patrick J. Mulrow*, Roberto Franco-Saenz* & Choh Hao Li†

* Department of Medicine, Medical College of Ohio, Toledo, Ohio 43699, USA
† Hormone Research Laboratory, University of California, San Francisco, California 94143, USA

Experimental and clinical evidence suggests that a non-ACTH pituitary factor(s) is important in the regulation of aldosterone secretion in certain conditions[1,2]. Earlier studies have suggested that β-melanotropin (β-MSH) has only weak steroidogenic potency on the adrenal gland in the rat[3,4], but recently Challis *et al.*[5] and Rudman *et al.*[6] have reported that α- and β-MSH have a trophic action on fetal adrenal gland. However, little is known about the effect of β-MSH on aldosterone production. We recently reported that β-lipotropin (β-LPH) stimulates aldosterone production in rat capsular cells[7]. As the amino acid sequence of β-MSH is contained within the β-LPH structure[8,9], we have now investigated the aldosterone-stimulating activity of synthetic β-MSH. Our results indicate that β-MSH causes a dose-dependent increase in aldosterone production in rat capsular cells which is blocked by a synthetic β-MSH analogue. The synthetic hormone does not increase corticosterone production in decapsular cells.

Collagenase-dispersed capsular (mainly zona glomerulosa cells) and decapsular (zona fasciculata and reticularis cells) adrenal cells from 20 female Sprague–Dawley rats on normal sodium diets were prepared for each experiment according to previously described methods[10]. In this capsular cell dispersion technique the degree of contamination with zona fasciculata cells has been reported to be less than 10% (ref. 11). Camel β-melanotropin (β_{Cl}-MSH) and the analogue, [Gly^{10}]-β_{Cl}-MSH, and ovine β-LPH were synthetic products as described previously[12–14].

As shown in Fig. 1*a*, 10^{-10} M [Asp^1, Ile^5]-angiotensin II (AII, Sigma) and ACTH$_{1-24}$ (Ciba) caused a significant increase in aldosterone production, maximal with 10^{-8} M AII or ACTH$_{1-24}$. Also, a significant increase in aldosterone was produced by 10^{-8} M β-MSH (from the control value 24.9 ± 3.6 (s.e.) ng per 100,000 capsular cells to 45.2 ± 3.9; $P < 0.01$) and maximum increase was obtained with 10^{-6} M β-MSH. The maximum aldosterone response obtained with β-MSH, however, is about the same as with AII. Figure 1*b* shows that ACTH$_{1-24}$ has a potent stimulating effect on corticosterone production by decapsular cells, whereas β-MSH has no effect. The effect of the β-MSH analogue on the response of capsular cells to half-maximal aldosterone-stimulating doses of β-MSH and β-LPH (3×10^{-8} M), and of ACTH$_{1-24}$ and AII (1.5×10^{-10} M) are shown in Fig. 2. The analogue caused ~30% inhibition of aldosterone production induced by β-MSH at a molar ratio of 100 : 1, and ~80% inhibition at a molar ratio of 1,000 : 1, and inhibited the β-LPH-induced aldosterone production to a similar extent at the same molar ratios. However, it did not affect ACTH$_{1-24}$ or AII-induced aldosterone production at molar ratios from 10 : 1 to 10,000 : 1. In one experiment we found that a 30,000 molar excess of β-MSH analogue did not affect ACTH-induced aldosterone production, and in another, that the β-MSH analogue itself has no effect on aldosterone production at doses from 10^{-9} M to 3×10^{-5} M.

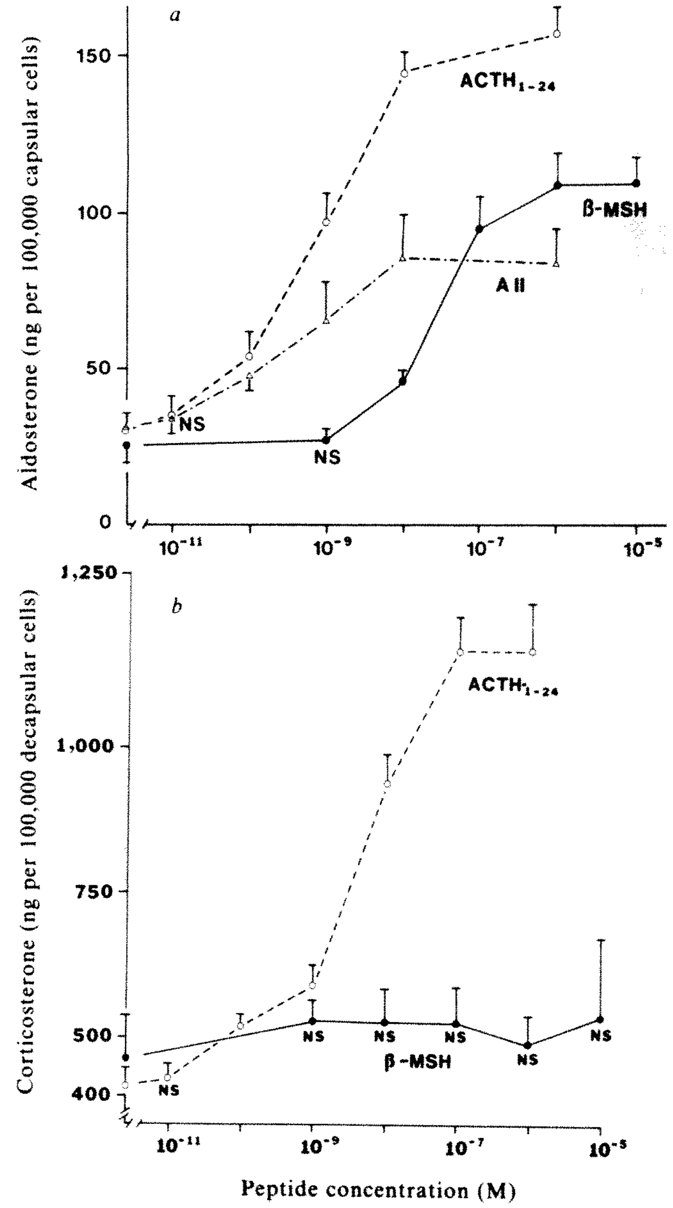


Fig. 1 *a*, Aldosterone production by AII, ACTH$_{1-24}$ and β-MSH in rat adrenal capsular cells. *b*, Corticosterone production by ACTH$_{1-24}$ and β-MSH in rat adrenal decapsular cells. Each point indicates mean ± s.e. from three different experiments. After decapitation, adrenal glands were removed and separated into capsular and decapsular portions. In *a*, capsular portions were minced into small pieces and incubated with collagenase (2 mg ml^{-1}) for 30 min. Collagenase-treated capsular cells were incubated in 1 ml medium 199 containing 2 mg ml^{-1} of bovine serum albumin in duplicate with various amounts of AII, ACTH$_{1-24}$ and β-MSH at 37 °C for 2 h. In *b*, collagenase-treated decapsular cells were incubated in duplicate with various amounts of ACTH$_{1-24}$ and β-MSH at 37 °C for 2 h. AII did not stimulate corticosterone production in decapsular cells (data not shown). NS: not significantly different from the control value; all other points significantly different from the control value ($P < 0.01$).

Studies by Page *et al.*[15] showed that α-MSH plus growth hormone stimulated aldosterone production in hypophysectomized sodium depleted rats. More recently, Vinson *et al.*[16]

---

‡ Present address: The 2nd Department of Internal Medicine, Faculty of Medicine, University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113, Japan.

0028-0836/81/200155—02$01.00

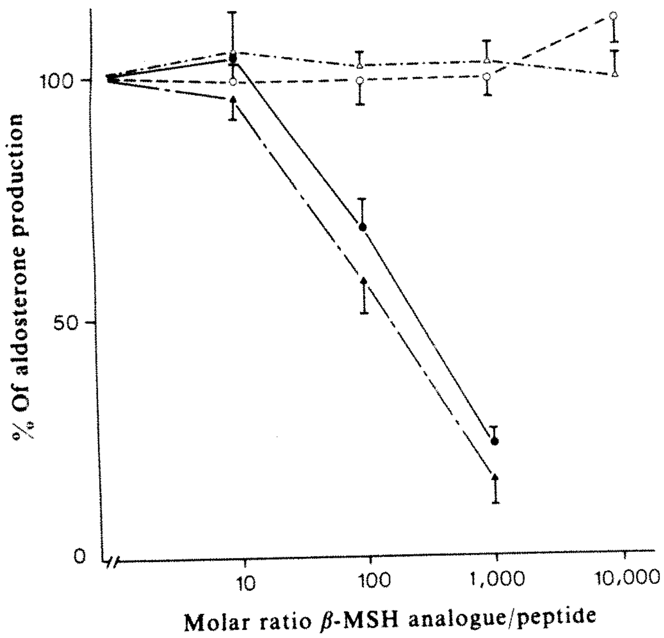


Fig. 2 Inhibition of aldosterone responses to half-maximal stimulating doses of β-MSH and β-LPH (3×10^{-8} M), and $ACTH_{1-24}$ and AII (1.5×10^{-10} M) by varying concentrations of β-MSH analogue. Collagenase-treated capsular cell suspensions were preincubated with various amounts (1.5×10^{-9} M to 3×10^{-5} M) of β-MSH analogue for 5 min and cell suspensions incubated with 3×10^{-8} M β-MSH or β-LPH, and with 1.5×10^{-10} M $ACTH_{1-24}$ or AII for 2 h. Each point shows mean ± s.e. from three different experiments. ●, β-MSH; ▲, β-LPH; ○, $ACTH_{1-24}$; △, AII.

reported that α-MSH has an adrenal zona glomerulosa-stimulating effect in the rat. The present data show that β-MSH has an aldosterone-stimulating effect on rat zona glomerulosa cells (Fig. 1*a*) and does not stimulate corticosterone production in rat adrenal decapsular cells (Fig. 1*b*). Although the concentration of β-MSH needed for half-maximal stimulation of aldosterone production was ~100 times greater than that of $ACTH_{1-24}$, we compared synthetic camel β-MSH with $ACTH_{1-24}$, which is known to be the biologically active sequence common to all species. As Lohmar *et al.*[17] and Lis *et al.*[18] reported some species differences in the lipolytic activity of β-LPH, it is possible that there are species differences in the aldosterone response to β-MSH. Moreover, recent evidence suggests that in humans β-MSH is an artefact due to degradation of β-LPH during extraction[19,20]. The physiological significance of β-MSH in the regulation of aldosterone secretion will therefore clearly require further study.

Because earlier studies showed that β-endorphin (β-LPH_{61-91}) has no steroidogenic properties[7], it seemed logical that γ-LPH (β-LPH_{1-58}) was the aldosterone-stimulating segment of β-LPH, and we have found that β-MSH (β-LPH_{41-58}), the carboxy-terminal segment of γ-LPH, contained aldosterone-stimulating activity. The demonstration that the β-MSH analogue inhibited both β-MSH and β-LPH activity to a similar extent suggests that the active core of β-LPH is β-MSH. Furthermore, the finding that the analogue inhibits β-LPH action but not $ACTH_{1-24}$ or AII actions suggests that the zona glomerulosa receptor for β-LPH or β-MSH may be different from the ACTH and AII receptors.

This work was supported in part by NIH grants HL-19644 and GM-2907.

Received 29 August 1980; accepted 18 February 1981.

1. Palmore, W. P. & Mulrow, P. J. *Science* **158,** 1482–1484 (1967).
2. McCaa, R. E., Young, D. B., Guyton, A. C. & McCaa, C. S. *Circulation Res.* Suppl. 34–35, 15 (1974).
3. Lowry, P. J., McMartin, C. & Peters, J. *J. Endocr.* **59,** 43–55 (1973).
4. Sayers, G., Swallow, R. L. & Giordano, N. D. *Endocrinology* **88,** 1063–1068 (1971).
5. Challis, J. R. G. & Torosis, J. D. *Nature* **269,** 818–819 (1977).
6. Rudman, D., Hollins, B. M., Lewis, N. C. & Chawla, R. K. *J. clin. Invest.* **65,** 822–828 (1980).
7. Matsuoka, H., Mulrow, P. J. & Li, C. H. *Science* **209,** 307–308 (1980).
8. Li, C. H., Barnafi, L., Chretien, M. & Chung, D. *Nature* **208,** 1093–1094 (1965).
9. Li, C. H. & Chung, D. Nature **260,** 623–624 (1976).
10. Matsuoka, H., Tan, S. Y. & Mulrow, P. J. *Prostaglandins* **19,** 291–298 (1980).
11. Tait, J. F., Tait, S. A. S., Gould, R. P. & Mee, H. R. J. *Proc. R. Soc.* **B185,** 375–407 (1974).
12. Li, C. H., Yamashiro, D. & Lemaire, S. *Biochemistry* **14,** 953–956 (1975).
13. Lemaire, S., Yamashiro, D., Rao, A. J. & Li, C. H. *J. med. Chem.* **20,** 155–158 (1977).
14. Yamashiro, D. & Li, C. H. *J. Am. chem. Soc.* **100,** 5174–5179 (1978).
15. Page, R. B., Boyd, J. E. & Mulrow, P. J. *Endocr. Res. Commun.* **1,** 53–62 (1974).
16. Vinson, G. P., Whitehouse, B. J., Dell, A., Etienne, T. & Morris, H. R. *Nature* **284,** 464–467 (1980).
17. Lohmar, P. & Li, C. H. *Endocrinology* **82,** 898–904 (1968).
18. Lis, M., Gilardeau, C. & Chretien, M. *Proc. Soc. exp. Biol. Med.* **139,** 680–683 (1972).
19. Bloomfield, G. A., Scott, A. P., Lowry, P. J., Gilkes, J. J. H. & Rees, L. H. *Nature* **252,** 492–493 (1974).
20. Gilkes, J. J. H. *et al. J. clin Endocr. Metab.* **40,** 450–457 (1975).

Circulating human pituitary pro-γ-melanotropin enhances the adrenal response to ACTH

Emad A. S. Al-Dujaili*‡, James Hope†, Fernando E. Estivariz†, Philip J. Lowry† & Christopher R. W. Edwards*

Departments of Endocrinology* and Chemical Pathology†,
St Bartholomew's Hospital, London EC1A 7BE, UK

The amino-terminal region of the common corticotropin/β-lipotropin (β-LPH) precursor has been identified in the AtT-20 mouse tumour cell as a glycopeptide with an apparent molecular weight of 16,000 (the '16K fragment')[1,2]. A third melanotropin core sequence or γ-MSH similar to that found in ACTH and β-LPH was predicted to occur in this glycopeptide from the complementary DNA sequence of mRNA isolated from bovine pituitary intermediate tissue[3]. Recently, the mouse 16K fragment has been found to have a small but significant potentiation on the corticosteroidogenesis elicited by ACTH in a static cell system, an effect that could be enhanced when the glycopeptide was pretreated with trypsin[4]. This synergism could also be mimicked by synthetic γ-MSH peptides *in vitro* and *in vivo*[5]. We report here the potentiating properties of a naturally occurring human pro-γ-MSH glycopeptide[6] on the ACTH-induced steroidogenic response of isolated perfused rat and human adrenocortical cells.

The *in vitro* superfusion studies of isolated adrenal cells used a modification of the system reported by Lowry and McMartin[7]. Female Wistar rats (200–220 g) were decapitated and the adrenal glands removed. The capsular tissue (namely zona glomerulosa cells) or the decapsulated tissue (zona fasciculata-reticularis cells) was then treated as previously described[8,9]. Two perfusion cell columns were prepared for each experiment; one column was always used as a control. The cells were checked for viability and counted before suspending in Bio-Gel. Each column contained $\sim1\times10^{6}$ cells. Using this technique the fasciculata-reticularis contamination of the glomerulosa preparation was 5–10% (ref. 10) as assessed by light microscopy and a Coulter counting system. For the preparation of isolated human adrenal cells, adrenocortical tissue obtained at total nephrectomy for renal carcinoma was treated in the same manner as described for rat adrenal tissue.

The cells were stimulated with $ACTH_{1-24}$ 1–5 ng l^{-1}, $ACTH_{1-39}$ 2–10 ng l^{-1}, Pro-γ-MSH_{1-77} (ref. 6) (γ-MSH precursor Trp 105 to Gln 29, numbering system of Nakanishi *et al.*[3]) 8×10^{-11}–5×10^{-9} mol l^{-1} (10-min pulses, 30–40 min between each pulse) alone or during γ-MSH precursor infusion and fractions were collected every 5 min. The concentrations of γ-MSH precursor used in these studies were based on the plasma levels of this polypeptide measured by radioimmunoassay[11]. The molecular size of the immunoactive γ-MSH

---

‡Present address: University Department of Medicine, Western General Hospital, Crewe Road, Edinburgh EH4 2XV, UK.

0028-0836/81/200156—04$01.00

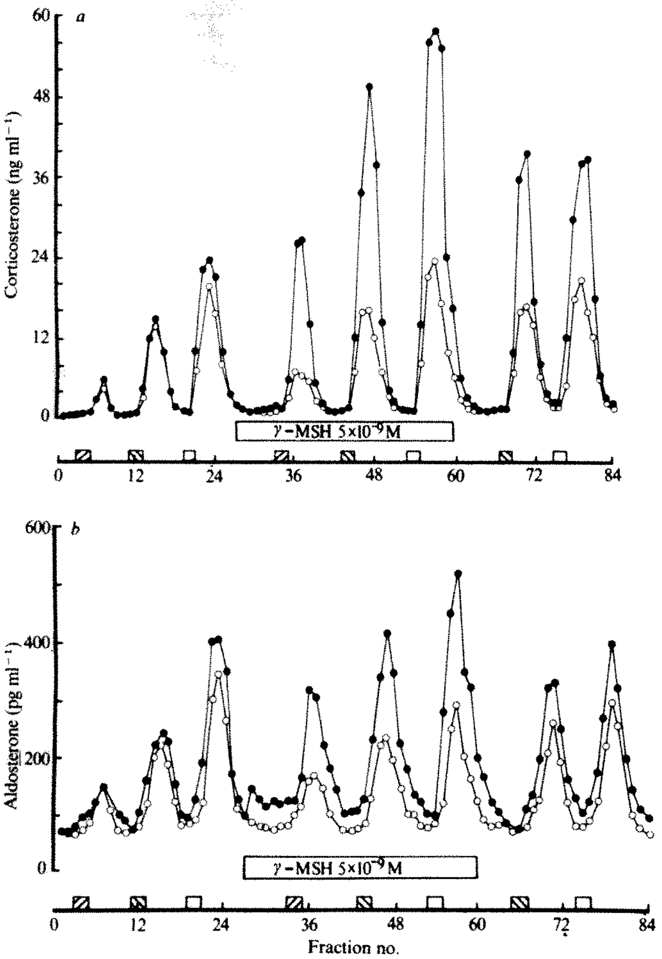


Fig. 1 Corticosterone (*a*) and aldosterone (*b*) secretion by isolated perfused rat zona glomerulosa cells in response to pulses of ACTH$_{1-24}$ given before, during and after an infusion of γ-MSH precursor (5×10^{-9} mol l^{-1}). The intact molecule of the precursor (77 amino acids) was found to circulate in the plasma of normal subjects and patients with ectopic ACTH syndrome, Nelson's syndrome and pituitary-dependent Cushing's disease as checked by affinity column chromatography. ACTH$_{1-24}$ and γ-MSH precursor were dissolved in 0.9% sodium chloride solution containing 10 mM HCl, 1 g l^{-1} bovine serum albumin, 1 g l^{-1} polypeptide (Sigma) and 100 mg l^{-1} lima bean trypsin inhibitor. Responses in the column infused with γ-MSH precursor (●) are compared with those in a control column (○). Doses of ACTH$_{1-24}$ given were 1 ng l^{-1} (▨), 2.5 ng l^{-1} (▧) and 5 ng l^{-1} (□). γ-MSH precursor was infused in the experimental columns only between fractions 27–60.

polypeptide in plasma corresponds to a 77-amino acid glycopeptide[11]. Aldosterone levels were measured in the superfusates by the specific, direct radioimmunoassay developed for human plasma[12], which as we have since shown[8,13] gives results comparable to those of methods which involve prior extraction of the aldosterone. Corticosterone was measured in the superfusates by a highly specific, direct radioimmunoassay[14], which has also been checked against techniques involving extraction[14]. The corticosterone assay uses ^{125}I-corticosterone–iodohistamine as ligand and highly specific corticosterone antiserum raised in rabbits against corticosterone-3-mono-oxime–bovine serum albumin immunogen. The specificity of the antibody was confirmed by the negligible cross-reactivities of this antiserum with potential interfering steroids.

ACTH$_{1-24}$ produced a dose-dependent increase in corticosterone and aldosterone production by rat zona glomerulosa cells (Fig. 1). When pro-γ-MSH$_{1-77}$ was infused in one column there was a significant potentiation of the corticosterone response to ACTH$_{1-24}$ which persisted even after the γ-MSH precursor infusion was stopped (Fig. 1*a*). The aldosterone response to ACTH was also potentiated by γ-MSH precursor but this potentiation was less marked than that observed for corticosterone (Fig. 1*b*). Similar results were obtained when rat zona glomerulosa cells were stimulated by ACTH$_{1-39}$ during γ-MSH precursor infusion. These findings were reproducible and Fig. 2 shows the responses of aldosterone and corticosterone produced by ACTH$_{1-24}$ and ACTH$_{1-39}$ using isolated superfused rat zona glomerulosa cells with or without the infusion of γ-MSH precursor (5×10^{-9} mol l^{-1}). There was a significant three- to fivefold increase ($P<0.001$) in corticosterone production and 0.75-fold ($P<0.01$) in aldosterone production by ACTH$_{1-24}$ and γ-MSH precursor infusion compared with ACTH$_{1-24}$ alone. ACTH$_{1-39}$ also produced a significant potentiation in corticosterone (1.5–2.5-fold, $P<0.001$) and aldosterone (1.0–2.0-fold, $P<0.001$) production during γ-MSH precursor infusion.

ACTH$_{1-24}$ and ACTH$_{1-39}$ produced a dose-dependent increase in corticosterone production by rat zona fasciculata-reticularis cells (Fig. 3*a*) which was also potentiated by continuous infusion of γ-MSH precursor, as can be seen by comparing the results in the experimental column with those produced by ACTH alone in the control column (Fig. 3*a*).

Pro-γ-MSH$_{1-77}$ alone (8×10^{-11}–4×10^{-9} mol l^{-1}) did not significantly increase steroidogenesis in either zona glomerulosa or zona fasciculata-reticularis cells, although a slight effect on aldosterone production was observed at the beginning of a continuous infusion (Fig. 1). Also, the precursor did not significantly potentiate the ACTH$_{1-24}$ response of cells of either tissue when given simultaneously with ACTH in bolus doses (as confirmed by Kruskal–Wallis one-way analysis of variance by ranks). The effect of an ACTH and γ-MSH precursor bolus dose on steroidogenesis was potentiated (Fig. 3*b*) following a 30-min priming dose of γ-MSH precursor (4×10^{-9} mol l^{-1}), and this effect could be maintained for some time by simultaneous bolus doses of the glycopeptide with ACTH. Isolated perfused human adrenal cells containing a mixture of zona glomerulosa and

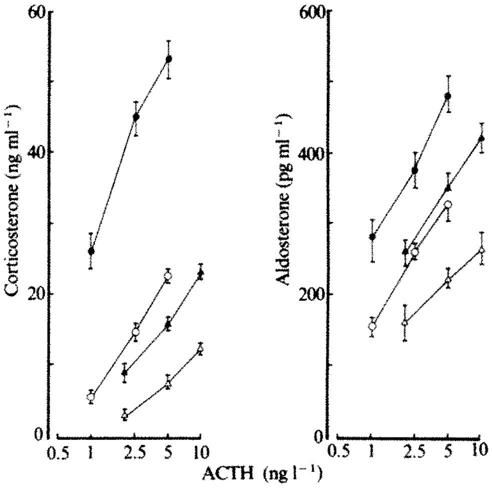


Fig. 2 A summary of the responses of aldosterone and corticosterone produced by ACTH$_{1-24}$ (1.0 ng l^{-1}, $n=5$; 2.5 ng l^{-1}, $n=5$; 5.0 ng l^{-1}, $n=6$) and ACTH$_{1-39}$ (2.0 ng l^{-1}, $n=6$; 5.0 ng l^{-1}, $n=4$; 10.0 ng l^{-1}, $n=5$) using isolated perfused rat zona glomerulosa cells with (●, ACTH$_{1-24}$; ▲, ACTH$_{1-39}$) or without (○, ACTH$_{1-24}$; △, ACTH$_{1-39}$) the infusion of pro-γ-MSH (5×10^{-9} mol l^{-1}). For corticosterone there was a three- to fivefold increase after ACTH$_{1-24}$ and pro-γ-MSH infusion compared with ACTH as bolus doses alone, which was significant ($P<0.001$) for all doses of ACTH$_{1-24}$. For aldosterone there was a 0.5–1.0-fold increase of the ACTH$_{1-24}$ and pro-γ-MSH infusion compared with ACTH alone. This potentiation was also significant ($P<0.01$ for ACTH 1.0 ng l^{-1}; $P<0.005$ for ACTH 2.5 ng l^{-1}; $P<0.001$ for ACTH 5.0 ng l^{-1}). For ACTH$_{1-39}$ the potentiation in corticosterone was 1.5–2.5-fold ($P<0.001$) and aldosterone 1.0–2.0-fold ($P<0.001$) during pro-γ-MSH infusion. Results are the mean ± s.e.m.

fasciculata-reticularis cells behaved similarly to rat adrenal cells, producing an increase in corticosterone (twofold) and aldosterone (20–40%) secretion to bolus doses of ACTH during an infusion of pro-γ-MSH (Table 1).

Thus, the intact pro-γ-MSH_{1-77} can potentiate the action of $ACTH_{1-24}$ or $ACTH_{1-39}$ on isolated superfused rat and human zona glomerulosa and zona fasciculata-reticularis cells to secrete corticosterone and aldosterone. This action was only apparent when the cells were exposed to a priming infusion or to a simultaneous continuous infusion and was more marked on corticosterone than on aldosterone production. It therefore seems that the potentiation of ACTH-induced steroidogenesis by the naturally occurrring intact form of pro-γ-MSH is time dependent. As there would be little or no degradation of pro-γ-MSH by contaminating peptidases in the superfused system used in our study, the physiological relationship to observations made with trypsin-treated 16K fragment[4] and synthetic γ-MSH[5] is unclear.

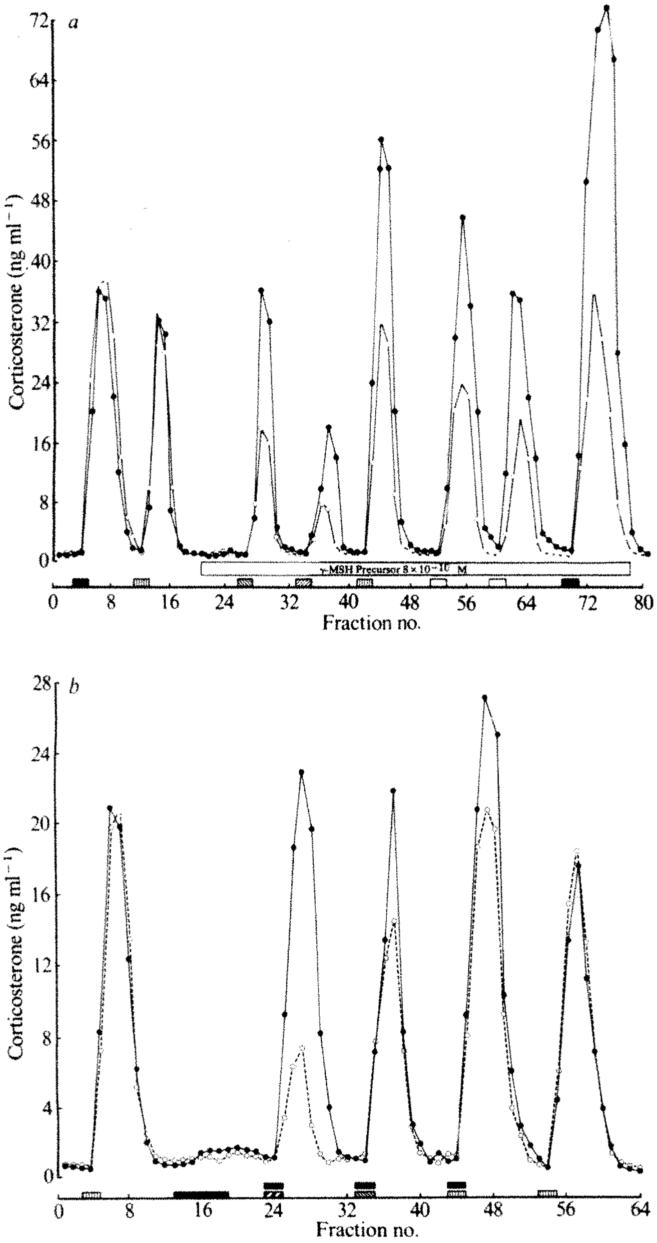


Fig. 3 *a*, Corticosterone secretion by isolated perfused rat zona fasciculata-reticularis cells in response to pulses of $ACTH_{1-24}$ and $ACTH_{1-39}$ given before, during and after an infusion of pro-γ-MSH (8×10^{-10} mol l^{-1}). Responses in the column infused with γ-MSH precursor (●) are compared with those in a control column (○). Doses of $ACTH_{1-24}$ given were 1 ng l^{-1} (☐), 2.5 ng l^{-1} (⬚) and 5 ng l^{-1} (■), and of $ACTH_{1-39}$, 2 ng l^{-1} (▨), 5 ng l^{-1} (▩) and 10 ng l^{-1} (▥). γ-MSH precursor was infused in the experimental column only between fractions 20–78. *b*, Corticosterone responses by isolated perfused rat zona glomerulosa cells in response to pulses of $ACTH_{1-24}$ (▨ 1 ng l^{-1}; ▩, 2.5 ng l^{-1}; ▥, 5 ng l^{-1}) given either alone in control column (○) or together with γ-MSH precursor (⬚, 4×10^{-9} mol l^{-1}) (●) after a 30-min priming infusion of γ-MSH precursor (⬚, fractions 13–19). The first and last bolus doses of ACTH in the experimental column were given without γ-MSH precursor.

Table 1 Aldosterone and corticosterone secretion by isolated perfused human adrenocortical cells in response to pulses of $ACTH_{1-24}$ (1–5 ng l^{-1}) given before and during an infusion of γ-MSH_{1-77}

Secretagogue		Aldosterone (pg ml^{-1}, mean ± s.d.)	Corticosterone (ng ml^{-1}, mean ± s.d.)
ACTH	Basal value	69.8 ± 9.0 (8)	0.5 ± 0.15 (8)
	ACTH (1 ng l^{-1})	135.0 ± 17.6 (5)	2.62 ± 0.45 (5)
	ACTH (2.5 ng l^{-1})	166.3 ± 17.8 (8)	3.28 ± 0.62 (8)
	ACTH (5 ng l^{-1})	183.3 ± 19.7 (5)	3.95 ± 0.92 (5)
ACTH + pro-γ-MSH (4×10^{-9} M)	Basal value	74.6 ± 11.6 (8)	0.55 ± 0.06 (8)
	ACTH (1 ng l^{-1})	187.5 ± 42.7* (4)	4.28 ± 0.83* (4)
	ACTH (2.5 ng l^{-1})	196.7 ± 5.8* (4)	5.74 ± 1.66* (4)
	ACTH (5 ng l^{-1})	250.0 ± 34.6* (4)	8.15 ± 2.98* (4)

* There was a significant potentiation ($P < 0.01$) in the response to ACTH during γ-MSH precursor infusion as confirmed by Mann-Whitney *U*-test. Values in parentheses represent the number of experiments.

We cannot altogether rule out the possibility that the potentiation of aldosterone production by zona glomerulosa cells is due to contaminating fasciculata cells providing the precursor corticosterone. This does occur in static systems but is less likely to occur in superfusion systems. Furthermore, the fact that ACTH and serotonin induce aldosterone secretion by our zona glomerulosa cells to a similar extent suggests that fasciculata contamination is not a major problem. We believe instead that the potentiation is related only to the effects of ACTH, a view confirmed by the lack of potentiating action of γ-MSH precursor on other zona glomerulosa stimulators such as angiotensin II, potassium or serotonin.

Our results suggest that the potentiating action of γ-MSH precursor may play an important part in modulating ACTH-induced steroidogenesis. A specific effect of two products of ACTH biosynthesis [α-MSH (ref. 15) and β-LPH (ref. 16)] on aldosterone production by zona glomerulosa cells has recently been observed and Pedersen *et al.*[5] have reported a direct effect of synthetic γ-MSH on adrenal cholesterol ester hydrolase activity. However, because, in our experiments γ-MSH precursor did not significantly induce steroidogenesis in the absence of ACTH it is uncertain whether the γ-MSH precursor has a direct steroidogenic action.

This work was supported by the MRC, the British Heart Foundation, the Joint Research Board, St Bartholomew's Hospital and the North East Thames Regional Research Scheme. We thank the surgeons in the Department of Urology, St Bartholomew's Hospital, for supplying human adrenal tissue and Dr B. C. Williams for collaboration in developing the direct radioimmunoassay for corticosterone.

Received 24 November 1980; accepted 23 February 1981.

1. Eipper, B. A. & Mains, R. E. *Endocr. Rev.* **1,** 1–27 (1980).
2. Eipper, B. A. & Mains, R. E. *J. supramolec. Struct.* **8,** 247–262 (1978).
3. Nakanishi, S. *et al. Nature* **278,** 423–427 (1979).
4. Pedersen, R. C. & Brownie, A. C. *Proc. natn. Acad. Sci. U.S.A.* **77,** 2239–2243 (1980).
5. Pedersen, R. C., Brownie, A. C. & Ling, N. *Science* **208,** 1044–1045 (1980).
6. Estivariz, F. E., Hope, J., McLean, C. & Lowry, P. J. *Biochem. J.* **191,** 125–132 (1980).
7. Lowry, P. J. & McMartin, C. *Biochem. J.* **142,** 287–294 (1974).
8. Al-Dujaili, E. A. S. thesis, Univ. London (1980).

9. Edwards, C. R. W. *et al. Clin. Endocr.* **13,** 45–50 (1980).
10. Tait, S. A. S., Tait, J. F., Gould, R. P., Brown, B. L. & Albano, J. D. *J. Steroid Biochem.* **5,** 775–787 (1974).
11. Hope, J., Ratter, S., Estivariz, F. E., McLoughlin, L. & Lowry, P. J. *Clin. Endocr.* (in the press).
12. Al-Dujaili, E. A. S. & Edwards, C. R. W. *J. clin. Endocr. Metab.* **46,** 105–113 (1978).
13. Al-Dujaili, E. A. S. & Edwards, C. R. W. *J. Steroid Biochem.* (in the press).
14. Al-Dujaili, E. A. S., Williams, B. C. & Edwards, C. R. W. *Steroids* **37,** 157–176 (1981).
15. Vinson, G. P., Whitehouse, B. J., Dell, A., Etienne, T. & Morris, H. R. *Nature* **284,** 464–467 (1980).
16. Matsuoka, H., Mulrow, P. J. & Li, C. H. *Science* **209,** 307–308 (1980).

Spleen colony-forming cell as common precursor for tissue mast cells and granulocytes

Y. Kitamura, M. Yokoyama, H. Matsuda & T. Ohno

Institute for Cancer Research and Department of Pathology,
Medical School, Osaka University, Kita-ku, Osaka, 530 Japan

K. J. Mori

Department of Microbiology, Faculty of Medicine, Kyoto University,
Sakyo-ku, Kyoto, 606 Japan

The haematopoietic stem cells which produce colonies in the spleen of irradiated mice (CFU-S)[1] can differentiate into erythrocytes, granulocytes, megakaryocytes and B lymphocytes[2–4]. Although mast cell precursors are known to be present in the bone marrow[5], spleen[6], fetal liver[7] and peripheral blood[8] of mice, the relationship between the mast cell precursor and CFU-S has remained unclear. We have now made use of mice of two mutant genotypes to determine whether or not the tissue mast cell is a progeny of CFU-S. Giant granules of beige (C57BL/6-bg^J/bg^J, Chediak-Higashi syndrome) mice can be used for identification of the origin of both tissue mast cells[5] and granulocytes[9], and WBB6F$_1$-W/W^v mice are useful recipients because they lack tissue mast cells owing to a defect in mast cell precursors[10]. We injected the cells from a single spleen colony into each WBB6F$_1$-W/W^v mouse and demonstrated directly that the tissue mast cell is a progeny of CFU-S.

Mice of strain WBB6F$_1$ (WB-$W/+$X C57BL/6-$W^v/+$)F$_1$-W/W^v and C57BL/6-bg^J/bg^J, +/+ were obtained from the Jackson Laboratory, Bar Harbor, Maine. As the haemoglobin pattern of WBB6F$_1$-W/W^v mice (diffuse, Hbb^d/Hbb^s) can easily be distinguished from that of C57BL/6 mice (sharp, Hbb^s/Hbb^s), it was used to identify the origin of erythrocytes[11]. The experimental design is shown in Fig. 1. X ray-irradiated (800 rad) C57BL/6-+/+ mice were injected with a mixture of bone marrow cells from beige and normal C57BL/6 mice. The irradiated C57BL/6 mice were killed 12 days after injection of bone marrow cells, spleen colonies were enucleated under the dissection microscope and cells from each colony suspended in serum from C57BL/6 mice. An aliquot of suspended cells was smeared, stained with Sudan Black B and counterstained with Giemsa solution to examine the proportion and type of granulocytes (beige or normal)[9]. The rest of the cells were injected into a WBB6F$_1$-W/W^v mouse intravenously. A blood sample was obtained from a retroorbital sinus 15 weeks after the cell injection. Haemoglobin pattern was determined by cellulose acetate electrophoresis after modification of haemoglobin with cystamine[12] and smears were stained with Sudan Black B to follow the appearance of beige-type granulocytes. After the examination of peripheral blood, the WBB6F$_1$-W/W^v recipient mice were killed and tissue distribution of mast cell types was determined (Table 1). The interval of 15 weeks between cell injection and killing of the recipients was chosen because previous experiments showed that the number of mast cells increased in tissues of WBB6F$_1$-W/W^v mice until the fifteenth week after the bone marrow transplantation[6,10].

Table 1 Type of mast cells in tissues of WBB6F$_1$-W/W^v mice 15 weeks after injection of a single spleen colony each

Tissues	No. of mice in which mast cells of each type appeared Beige type only	Normal type only	Both types	Total
Skin	0	0	0	0
Mesentery	0	2	0	2
Caecum	2	4	0	6
Forestomach	2	6	2	10
Glandular stomach	3	7	2	12
Spleen	10	7	3	20
Whole body*	10	9	4	23

Tissues of 106 WBB6F$_1$-W/W^v mice were examined. Spleen, stomach, caecum and a piece of dorsal skin were removed and fixed in 10% buffered formalin (*p*H 7.2). The stretched preparation of the mesentery was also fixed in 10% formalin. Tissues were embedded in paraffin; sections (5 μm thick) and stretched mesentery were stained with acidified toluidene blue (*p*H 3.0). Criteria to determine the appearance of mast cells in the skin, mesentery, forestomach and glandular stomach are described elsewhere[6]. Mast cells were considered to have developed in the spleen when more than 20 mast cells were detected per longitudinal section.

* No. of mice in which mast cells appeared in more than one of above tissues.

A total of 117 discrete colonies were enucleated from the spleen possessing 1–3 colonies which developed after injecting the mixture of 10^4 beige and 10^4 normal bone marrow cells. In 36 out of 117 colonies, more than 10% of cells in one colony were granulocytes. Beige-type granulocytes with giant sudanophilic granules were predominant (>90%) in 19 out of 36 colonies, and normal-type granulocytes predominated in 16 colonies. A considerable number of granulocytes of both types were observed in only 1 of 36 colonies (15% beige type, 85% normal type). Thus, each spleen colony could easily be identified as from either a beige C57BL/6 mouse or a normal C57BL/6 mouse.

One hundred and six WBB6F$_1$-W/W^v mice which each received cells from a single spleen colony survived for 15 weeks. A change of haemoglobin pattern from diffuse to sharp was detected in 3 out of 106 WBB6F$_1$-W/W^v mice. Only 1 of 106 WBB6F$_1$-W/W^v mice showed beige-type granulocytes (3%) in the peripheral blood. As shown in Table 1, mast cells appeared in more than one of the tissues examined in 23 out of 106 WBB6F$_1$-W/W^v mice. All the mast cells scored were of either beige or normal type in 19 out of 23 mice. The type of granulocytes in the original spleen colony had been determined in 10 of 19 such cases. The relationship between the type of granulocytes and that of the mast cells is shown in Table 2. The type of the

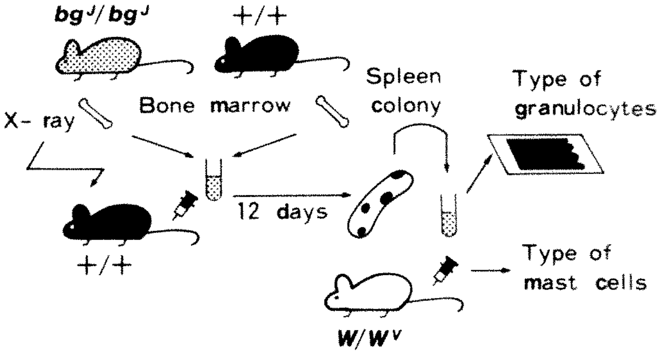


Fig. 1 Experimental design for comparison of the type of granulocytes in a spleen colony with that of tissue mast cells which appeared in WBB6F$_1$-W/W^v mouse after injection of cells from each colony.

0028–0836/81/200159—02$01.00

Table 2 Comparison of the type of granulocytes in a spleen colony with that of mast cells which developed in tissues of a WBB6F$_1$-W/W^v mouse after injection of cells from each colony

Mouse no.	Type of granulocytes (beige : normal)	Type of mast cells (beige : normal) Spleen	Glandular stomach	Forestomach	Caecum
2410	100:0	152:0	ND	ND	ND
2418	92:8	198:0	ND	ND	ND
2428*	6:94	0:317	0:253	0:48	0:383
3011	100:0	111:0	ND	ND	ND
3189	3:97	ND	0:354	ND	ND
3638	100:0	136:0	ND	ND	ND
4141*	100:0	373:9	64:0	120:0	38:0
4145	100:0	615:0	ND	ND	ND
4149	0:100	0:650	ND	ND	ND
4153	0:100	0:87	ND	0:381	ND

Mast cells did not develop in the mesentery and skin of all mice listed above. ND, mast cells did not develop.

* Haemoglobin pattern changed from diffuse to sharp in mice 2428 and 4141. Beige-type granulocytes (3%) appeared in the peripheral blood of mouse 4141.

Table 3 Development of mast cells in WBB6F$_1$-W/W^v mice 15 weeks after injection of various cells

Group no.	Origin of injected cells	No. of mice examined	No. of mice in which mast cells of each type appeared* Beige type only	Normal type only	Both types	Total
1	Spleen fragment†‡ without colony	18	0	0	0	0¶
2	Colony from spleen‡§ with 1–3 colonies	106	10	9	4	23
3	Colony from spleen‖ with 7–11 colonies	69	6	5	18#	29**

* Mast cells appeared in more than one of tissues listed in Table 1.

† Areas of spleen other than colonies were excised from the spleen with 1–3 colonies.

‡ The mixture of bone marrow cells from beige (10^4) and normal (10^4) C57BL/6 mice was injected to make colonies.

§ Data from Table 1.

‖ The mixture of bone marrow cells from beige (2×10^4) and normal (2×10^4) C57BL/6 mice were injected to make colonies.

¶ $P<0.05$ by χ^2 test, when the frequency of mice in which mast cells appeared was compared between groups 1 and 2.

# $P<0.01$, when the frequency of mice in which both types of mast cells appeared was compared between groups 2 and 3.

** $P<0.01$, when the frequency of mice in which mast cells appeared was compared between groups 2 and 3.

granulocytes and that of the tissue mast cells were exclusively consistent in all cases.

In the above experiment, mast cells appeared in only 23 of 106 WBB6F$_1$-W/W^v mice after injecting cells from an individual spleen colony. Two experiments were carried out to obtain more information on this low acceptance of spleen-colony cells. First, as there is a possibility that colonies were not excised precisely from the spleen, cells from the other areas of the spleen (over a diameter of about 2 mm) rather than the colony were injected to WBB6F$_1$-W/W^v mice. No mast cells appeared in the examined tissues of the W/W^v recipients 15 weeks after the cell injection (Table 3). Second, in the experiment shown in Tables 1 and 2, discrete colonies were enucleated from the spleen possessing 1–3 colonies to minimize contamination of cells derived from more than two colonies. However, the number of mast cell precursors per colony might be dependent on the number of colonies per spleen because there are controversial reports about the effect of graft size on proliferation and differentiation of CFU-S[13,14]. Although the efficiency of mast cell development increased after injection of an individual colony enucleated from the spleen possessing 7–11 colonies, both beige and normal types of mast cells appeared in more than half cases (Table 3). The latter result illustrates the difficulty of enucleating an individual colony from a spleen with 7–11 colonies, so the increased efficiency of mast cell development could not simply be attributed to the effect of graft size.

The present results directly demonstrate that precursors of tissue mast cells are generated within a spleen colony produced by CFU-S. Because the type of granulocytes in the colony matched the type of mast cells in tissues of WBB6F$_1$-W/W^v mice, CFU-S seems to be a common precursor for granulocytes and tissue mast cells.

Although a considerable proportion of granulocytes was observed in only 36 out of 117 spleen colonies examined, most of the colonies contained a considerable proportion of erythroid cells. Since Wu *et al.*[2] have shown that erythroid cells and granulocytes in the individual spleen colony are derived from a single CFU-S, at least a portion of CFU-S seems to have the potential to differentiate into erythrocytes, granulocytes and tissue mast cells. In fact, after injection of a beige-type spleen colony into a WBB6F$_1$-W/W^v mouse (4141 in Table 2), haemoglobin pattern changed from diffuse to sharp, beige-type granulocytes appeared in the peripheral blood, and beige-type mast cells developed in the spleen, stomach and caecum.

We thank K. Matsumoto for valuable discussions. This work was supported in part by the Japanese Ministry of Education, Science and Culture, and Mitsuhisa Memorial Foundation.

Received 7 November 1980; accepted 9 March 1981.

1. Till, J. E. & McCulloch, E. A. *Radiat. Res.* **14,** 213–222 (1961).
2. Wu, A. M., Till, J. E., Siminovitch, L. & McCulloch, E. A. *J. Cell Physiol.* **69,** 177–184 (1967).
3. Curry, J. L. & Trentin, J. J. *Devl Biol.* **15,** 395–413 (1967).
4. Lala, P. K. & Johnson, G. R. *J. exp. Med.* **148,** 1468–1477 (1978).
5. Kitamura, Y., Shimada, M., Hatanaka, K. & Miyano, Y. *Nature* **268,** 442–443 (1977).
6. Kitamura, Y. *et al. J. exp. Med.* **150,** 482–490 (1979).
7. Kitamura, Y., Shimada, M. & Go, S. *Devl Biol.* **70,** 510–514 (1979).
8. Kitamura, Y., Hatanaka, K., Murakami, M. & Shibata, H. *Blood* **53,** 1085–1088 (1979).
9. Murphy, E. D., Harrison, D. E. & Roth, J. B. *Transplantation* **15,** 526–530 (1973).
10. Kitamura, Y., Go, S. & Hatanaka, K. *Blood* **52,** 447–452 (1978).
11. Harrison, D. E., Astle, C. M. & DeLaittre, J. A. *Blood* **54,** 1152–1157 (1979).
12. Whitney, J. B. III *Biochem. Genet.* **16,** 667–672 (1978).
13. Schofield, R. & Pozzi, L. V. in *Advances in Transplantation. Proceedings of the First International Congress of the Transplantation Society* (eds Dausset, J., Hamburger, J. & Mathé, G.) 423–425 (Munksgaard, Copenhagen, 1969).
14. Metcalf, D. & Moore, M. A. S. in *Hematopoietic Cells* 90–92 (North-Holland, Amsterdam, 1971).

Immunochemical evidence for hybrid sialoglycoproteins of human erythrocytes

W. J. Mawby*, D. J. Anstee† & M. J. A. Tanner*

* Department of Biochemistry, University of Bristol,
Bristol BS8 1TD, UK
† South Western Regional Blood Transfusion Centre, Southmead,
Bristol BS10 5ND, UK

The two major sialoglycoproteins of the human erythrocyte membrane (α and δ, glycophorins A and B[1,2]) have identical amino acid sequences for the first 26 residues from the amino terminus, except that α expresses M or N blood group antigen activity whereas δ carries only blood group N activity. In addition, the asparagine at position 26 on α carries an oligosaccharide chain which is absent from the same position on δ. The two sialoglycoproteins differ in their remaining amino acid sequence and δ expresses blood group Ss activity. There are also variant sialoglycoproteins which have properties of both the α and δ molecules[1,3] and may be hybrids of these. Using antibodies directed against different structural regions of the major sialoglycoprotein α, we confirm here that two variant erythrocytes (Miltenberger class V (MiV) and Ph) contain hybrid sialoglycoprotein molecules (Fig. 1). These hybrid sialoglycoproteins arise from cross-over events between the genes coding for α and δ. It is suggested that the two genes are closely associated in the order α, δ ($5' \rightarrow 3'$) on the chromosome.

The two antibodies used in this study were R18 and I16. R18 is specific for a portion of the extracellular region of α and I16 for the intracellular portion (and/or intramembranous portion; see Fig. 1). The R18 antibody (LICR.LON/R18) is a mouse monoclonal antibody that reacts with untreated and trypsin-treated normal erythrocytes but does not react with En(a−) erythrocytes which lack α (ref. 4). Trypsin removes part of the extracellular region of α, thus the binding site of the antibody lies between the extracellular residues 40 and 72 (ref. 4). I16 was produced in a rabbit by repeated subcutaneous injections of partially purified sialoglycoprotein α in Freund's complete adjuvant. The antigen was partially purified by butanol extraction of erythrocyte membranes[5], followed by chromatography on Sepharose 6B equilibrated in 5 mM sodium phosphate *p*H 8.0, 0.1% Ammonyx LO[6]. The rabbit was bled 7–10 days after the fourth and subsequent injections. The antibody was shown to be specific for α by two-dimensional immunoelectrophoresis and immunoprecipitation using detergent-solubilized radioiodinated erythrocytes. To remove antibodies specific for the extracellular portion of α, the antiserum was repeatedly absorbed with intact erythrocytes until no haemagglutination activity remained, and the amount of solubilized α immunoprecipitated remained unchanged. This antibody therefore recognizes determinants between residues 73 and 131 of α (ref. 7).

The two antibodies were used to immunoprecipitate membrane proteins of normal and variant human erythrocytes which had been labelled with ^{125}I using lactoperoxidase. The variant erythrocytes studied were MiV erythrocytes from both homozygous[8] and heterozygous individuals[1], and erythrocytes heterozygous for the Ph condition[3]. These variants contain abnormal sialoglycoproteins designated $(\alpha\text{–}\delta)^{\mathrm{MiV}}$ in MiV erythrocytes[1] and $(\delta\text{–}\alpha)^{\mathrm{Ph}}$ in Ph erythrocytes[3]. Figure 2 (*f*, *g*, *i*, *k*) shows that the $(\alpha\text{–}\delta)^{\mathrm{MiV}}$ component is immunoprecipitated by R18 but not by I16. This sialoglycoprotein therefore contains an antigenic determinant of α in the extracellular region between residues 40 and 72, but lacks the determinant recognized by I16 in the remaining C-terminal portion. Previous evidence suggested that $(\alpha\text{–}\delta)^{\mathrm{MiV}}$ has features of both α and δ components[1,8] as it carries weak blood group M, N antigen activity and has the characteristic trypsin sensitivity of normal α, but also carries blood group s activity which is associated with α[9,10]. The present results therefore provide independent confirmation that the variant sialoglycoprotein of MiV erythrocytes is a hybrid molecule consisting of the amino-terminal portion of α, and the carboxy-terminal portion of δ, or more accurately an unusual allele of δ[1,8].

Figure 2 (*h*, *l*) shows that the $(\delta\text{–}\alpha)^{\mathrm{Ph}}$ component of Ph erythrocytes is immunoprecipitated by I16 but not by R18. The $(\delta\text{–}\alpha)^{\mathrm{Ph}}$ sialoglycoprotein therefore contains an antigenic determinant of α in the intracellular or intramembranous regions, but does not contain the extracellular determinant recognized by R18. We have previously suggested that the $(\delta\text{–}\alpha)^{\mathrm{Ph}}$ component contains the amino-terminal portion of δ as it is trypsin resistant, lacks an *N*-glycosidically linked oligosaccharide and blood group M activity, but carries blood group N activity[3]. The present data confirm that the $(\delta\text{–}\alpha)^{\mathrm{Ph}}$ component has an amino-terminal portion which differs from that of normal α, and conclusively shows that the component contains a determinant of the carboxy-terminal region of α. This provides strong evidence that the $(\delta\text{–}\alpha)^{\mathrm{Ph}}$ component is a hybrid molecule consisting of the amino-terminal portion of δ and the carboxy-terminal portion of α.

We suggest that the hybrid molecules arise by a mechanism similar to that which produces the abnormal Lepore-type haemoglobins[11,12], which occur when misalignment of genes leads to unequal crossing-over and generates new genes with characteristics of both the parent genes. The misalignment of genes coding for α and δ may result from the sequence

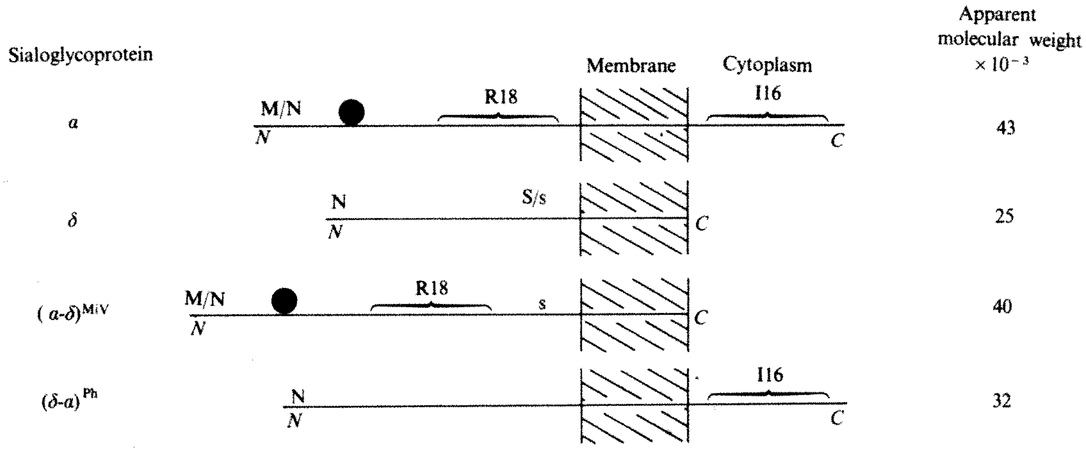


Fig. 1 Schematic diagram of the structures of normal and variant sialoglycoproteins. *N* and *C* represent the amino and carboxy termini of each sialoglycoprotein; M, N, S, s represent the blood group antigens associated with each molecule. The filled circle shows the approximate position of the asparagine-linked oligosaccharide when present. This oligosaccharide has been shown to be present on $(\alpha\text{–}\delta)^{\mathrm{MiV}}$ using a lectin binding technique[3] with the *Phaseolus vulgaris* lectin (phytohaemagglutinin, data not shown). Also shown are the approximate locations of the antigenic sites recognized by the antibodies R18 and I16. The apparent molecular weights are derived from the electrophoretic mobility of the proteins on SDS-polyacrylamide gels run in the same conditions as those described in Fig. 2 legend.

0028-0836/81/200161—02$01.00

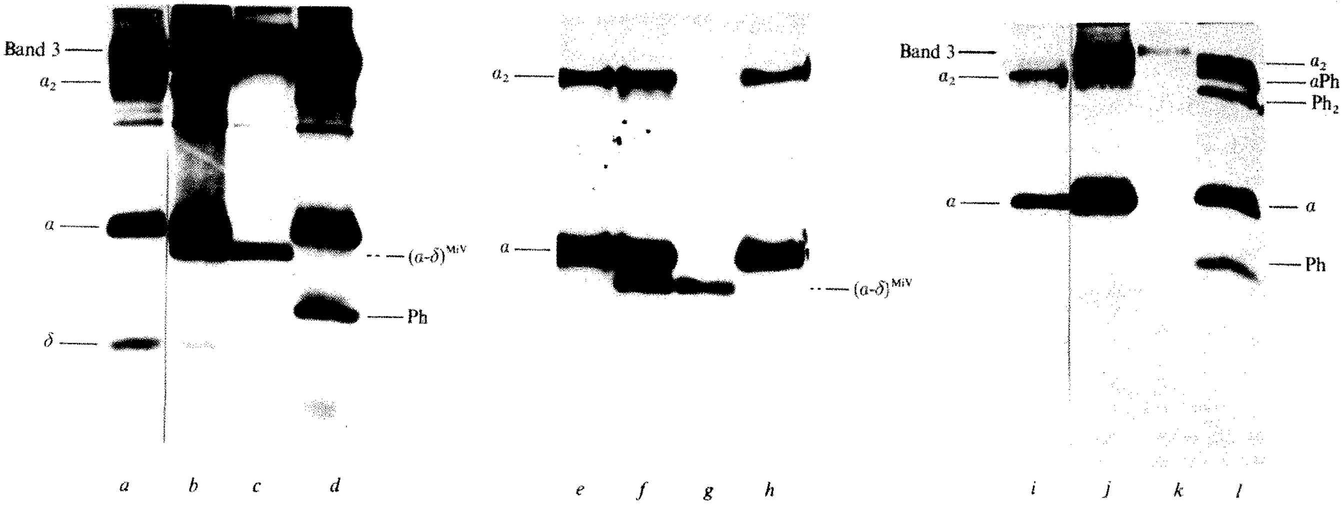


Fig. 2 Immunoprecipitation of labelled sialoglycoproteins from variant erythrocytes. Intact erythrocytes from donor M.P. (Ph variant), donor L.R. (heterozygous for the MiV phenotype), donor F.M. (homozygous for the MiV phenotype) and a normal donor were radioiodinated using lactoperoxidase[1]. The cells were washed and solubilized in 4 vol of 150 mM NaCl, 50 mM Tris-HCl pH 8.0, 1% w/v Na deoxycholate[16]. When using R18 the solubilized sample was initially treated with mouse non-immune serum for 1 h, followed by rabbit anti-mouse IgG for 30 min and 200 μl of a 10% suspension of fixed *Staphylococcus aureus* (strain Cowan I) for a further 30 min, all at 0 °C. The sample was then centrifuged and 10 μl of R18 antibody was added to the supernatant, followed by 10 μl of rabbit anti-mouse IgG (1 mg ml^{-1}) and 50 μl of 50% Protein-A Sepharose (Pharmacia); times and temperature as for the initial treatment. The mixture was centrifuged and the pellet washed three times with 0.15 M NaCl, 5 mM EDTA, 0.5% Triton X-100, all at 0 °C. The pellet was then suspended in 20 mM Tris pH 8.0, 5 mM EDTA, 5% SDS, 10% glycerol, 0.1 mg ml^{-1} bromophenol blue, 2% β-mercaptoethanol, boiled for 3 min, centrifuged and the supernatant electrophoresed in a 10% slab gel with a 5% overlay, using the buffer system of Laemmli[17]. When using I16 antibody the sample was first treated with normal rabbit IgG and *S. aureus* as above. The supernatant was immunoprecipitated with 10 μl of I16 followed by 50 μl protein-A Sepharose and the pellet treated as above. The SDS-polyacrylamide gels were run overnight and fixed in methanol/acetic acid/water (5:1:5) for 30 min, washed, dried and exposed to Kodak No-screen film for 1–5 days. *a–d*, Autoradiographs of radioiodinated intact erythrocytes before immunoprecipitation. *a*, Control erythrocytes; *b*, MiV heterozygous erythrocytes; *c*, MiV homozygous erythrocytes and *d*, Ph erythrocytes. *e–h*, Autoradiograph of ^{125}I-labelled components immunoprecipitated using R18. *e*, Control erythrocytes; *f*, MiV heterozygous erythrocytes; *g*, MiV homozygous erythrocytes and *h*, Ph erythrocytes. *i–l*, Autoradiograph of ^{125}I-labelled components immunoprecipitated using an IgG fraction of I16 antibody. *i*, Control erythrocytes; *j*, MiV heterozygous erythrocytes; *k*, MiV homozygous erythrocytes and *l*, Ph erythrocytes. The apparent molecular weights of the monomer bands are given in Fig. 1. α_2, αPh and Ph_2 represent the dimer of α, the complex of α with Ph and the dimer of Ph respectively. The band in lane *k* represents nonspecific background precipitation of band 3.

homology of the first 26 residues of the two proteins[9,10,13]. It follows from this interpretation of the origin of the hybrid sialoglycoproteins that the genes coding for α and δ are closely linked on the same chromosome. Furthermore, the gene coding for α must be on the 5′ side of that coding for δ as the event which gives rise to the α–δ hybrid sialoglycoprotein ($(\alpha\text{–}\delta)^{\mathrm{MiV}}$) also results in the absence of both normal α and δ[1,8], a situation similar to that occurring in Lepore-type haemoglobins. Conversely, the event which gives rise to a δ–α hybrid would not exclude normal α and δ, a situation analogous to the anti-Lepore haemoglobins. The $(\delta\text{–}\alpha)^{\mathrm{Ph}}$ component, a δ–α hybrid sialoglycoprotein, is associated with the presence of α but not δ[3]. This absence of δ is unexpected and we have attributed it to a cross-over event between a normal chromosome containing α and δ and one lacking δ. The latter are common in the black African population from which the variant originated[14].

This work was supported in part by a grant from the MRC. We thank Dr P. A. W. Edwards for the monoclonal antibody LICR/LON. R18.

Note added in proof: The genes giving rise to the MN and Ss antigens and thus α and δ have recently been assigned to the *q*28 or *q*31 region of chromosome 4 (ref. 15).

Received 2 January; accepted 17 February 1981.

1. Anstee, D. J., Mawby, W. J. & Tanner, M. J. A. *Biochem. J.* **183,** 193–203 (1979).
2. Furthmayr, H. *J. supramolec. Struct.* **9,** 79–95 (1978).
3. Tanner, M. J. A., Anstee, D. J. & Mawby, W. J. *Biochem. J.* **187,** 493–500 (1980).
4. Anstee, D. J. & Edwards, P. A. W. (in preparation).
5. Anstee, D. J. & Tanner, M. J. A. *Biochem. J.* **138,** 381–386 (1974).
6. Furthmayr, H., Tomita, M. & Marchesi, V. T. *Biochem. biophys. Res. Commun.* **65,** 113–121 (1975).
7. Tomita, M., Furthmayr, H. & Marchesi, V. T. *Biochemistry* **17,** 4756–4770 (1978).
8. Vengelen-Tyler, V. *et al. Transfusion* **21,** 1 (1981).
9. Dahr, W., Gielen, W., Beyreuther, K. & Kruger, J. *Hoppe-Seyler's Z. physiol. Chem.* **361,** 145–152 (1980).
10. Dahr, W., Beyreuther, K., Steinbach, H., Gielen, W. & Kruger, J. *Hoppe-Seyler's Z. physiol. Chem.* **361,** 895–906 (1980).
11. Bunn, H. F., Forget, B. G. & Ranney, H. M. *Human Haemoglobins*, 151 (Saunders, Philadelphia, 1977).
12. Weatherall, D. J. & Clegg, J. B. *Cell* **16,** 467–479 (1977).
13. Furthmayr, H. *Nature* **271,** 519–524 (1978).
14. Lowe, R. F. & Moores, P. *Hum. Hered.* **22,** 344 (1972).
15. Cook, P. J. L., Noades, J. E., Lomas, C. G., Buckton, K. E. & Robson, E. B. *Ann. hum. Genet.* **44,** 61 (1980).
16. Edwards, P. A. W. *Biochem. Soc. Trans.* **8,** 334 (1980).
17. Laemmli, U. K. *Nature* **227,** 680 (1970).

Properties of postsynaptic channels induced by acetylcholine in different frog muscle fibres

R. Miledi & O. D. Uchitel

Department of Biophysics, University College London,
Gower Street, London WC1E 6BT, UK

Skeletal muscles in the frog are composed of two distinct classes of muscle fibre: fast muscle fibres capable of propagating action potentials and twitches, and slow muscle fibres normally unable to generate action potentials or twitches[1,2]. In addition, amphibian muscles contain a spectrum of 'intermediate' fibres whose structural and functional properties lie between those of fast and slow fibres[3–6]. Much is now known about the characteristics of the channels opened by the transmitter acetylcholine (ACh) acting on the membrane of fast fibres[7,8], but the molecular action of ACh on the other fibre types is only poorly understood. We report here the existence of a muscle in the mandibular arch of the frog in which most, if not all, the fibres are multiply innervated and are capable of eliciting action potentials. We also report that the channels induced by the transmitter on the synaptic membrane of fast, slow and submaxillaris muscle fibres differ in their lifetimes and conductances.

0028-0836/81/200162—04$01.00

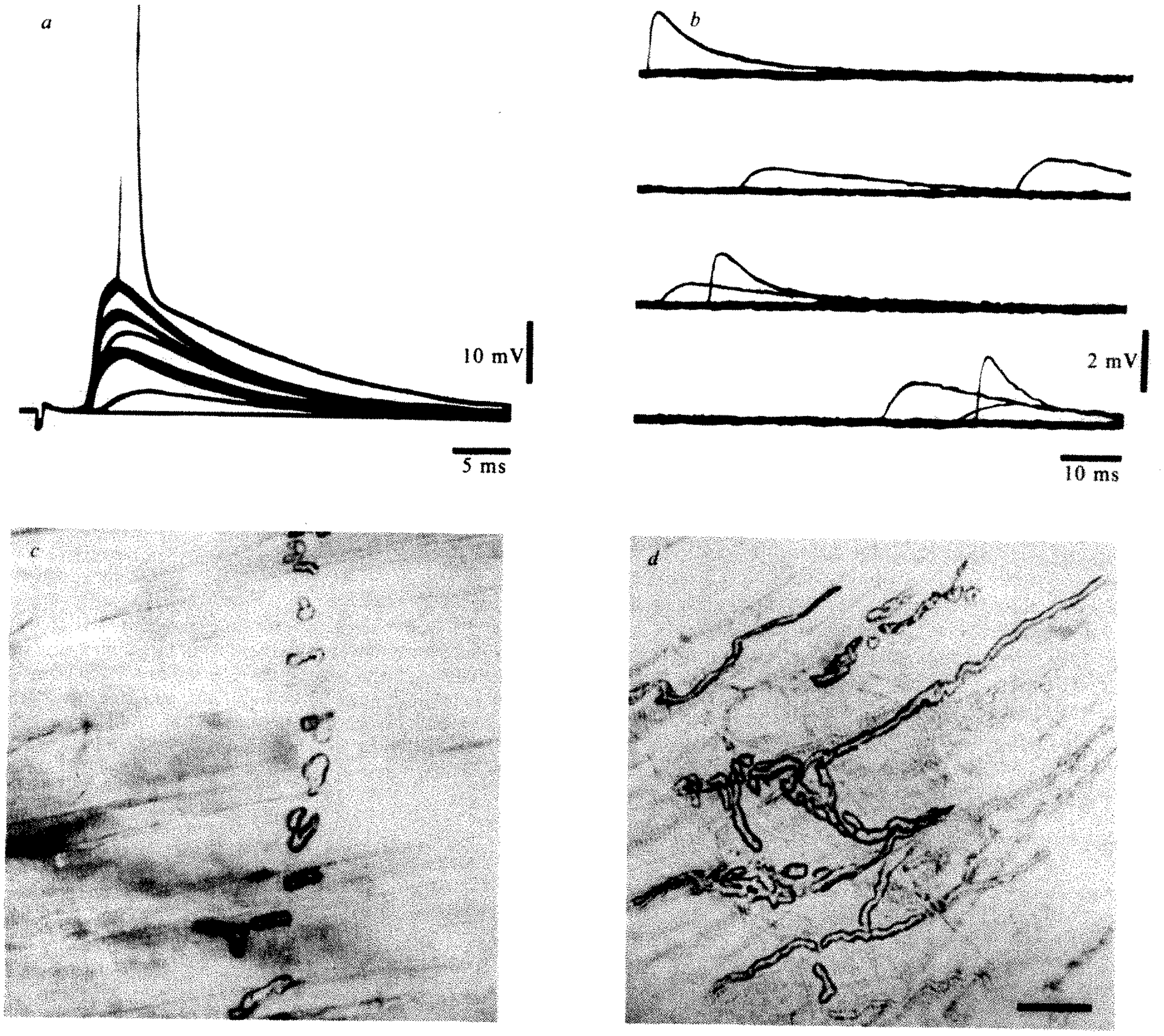


Fig. 1 Endplate potentials recorded intracellularly from a fibre in a submaxillaris muscle in Ringer with 5.4 mM Ca^{2+}. The muscle was preincubated in 2 M formamide for 90 min to block the contractile responses[23]. *a*, The superimposed records show five components obtained by increasing progressively the strength of the stimulus applied to the submaxillaris nerve. An action potential was elicited by the compound endplate potential (Resting potential, 85 mV, temperature, 22 °C.) *b*, Miniature endplate potentials recorded from a fibre in the same muscle as in *a*. *c*, *d*, Partial views of two different regions of the same submaxillaris muscle where cholinesterase was stained with Karnovsky and Roots method[24]. Scale bar, 20 μm.

Experiments were performed on fast and slow muscle fibres of the pyriformis muscle[9], and on the fibres of the submaxillaris muscle[9] of the frog *Rana temporaria*. The muscles, with their nerve supply, were isolated and bathed in normal Ringer solution (in mM: NaCl 115, KCl 2.5, $CaCl_2$ 1.8; buffered with HEPES at *p*H 7.3), or in low Ca^{2+}–high Mg^{2+} Ringer (same as normal, but $CaCl_2$ 0.3 mM and $MgCl_2$ 5 mM), kept at about 6 °C. The different types of muscle fibres were identified by one or more of the following characteristics: passive and active electrical membrane properties, innervation pattern and contractile behaviour[1–6,10,11].

While surveying the ACh sensitivity and distribution of cholinesterase in frog muscle, Couteaux and Miledi (unpublished) found that in the submaxillaris muscle the fibres were innervated at multiple sites and were capable of propagating action potentials. We decided to examine the muscle in more detail and, while the work was in progress, Kordylewski[5,6] reported a morphological study of this muscle, which he calls submandibular.

Muscle fibres in the submaxillaris muscle are 8–16 mm long and extend from a narrow central tendon to the jaw. They are relatively small (5–50 μm diameter) and even after impalement with two low-resistance microelectrodes their input resistance is several MΩ. All the fibres examined are able to generate an action potential when stimulated, directly with an intracellular microelectrode or indirectly through its motor nerve (Fig. 1*a*), and the action potential triggers a phasic contraction. All the muscle fibres are innervated by the submaxillaris nerve, but the posterior part of the muscle is also innervated by the subhyoideus nerve. When the submaxillaris nerve is stimulated, alone or together with the subhyoideus nerve, with stimuli of increasing intensity, the endplate potentials generated are made up of several discrete components (Fig. 1*a*), indicating that individual muscle fibres are innervated by several axons. Miniature endplate potentials are of low frequency and can be recorded anywhere along the muscle fibres down to their tendon ends. Their amplitude and time course are very variable (Fig. 1*b*), as would correspond to local events originating at various distances from the recording microelectrode.

As expected from the electrophysiological experiments, histochemical staining revealed synaptic contacts all over the muscle. Endplates vary in shape from simple small bouton-like contacts to long branched ones (Fig. 1*c*,*d*). Individual muscle fibres have 5–10 endplates distributed along their length, and both bouton and branched types seem to coexist on the same fibre.

Table 1 Characteristics of the synaptic ACh-activated channels obtained from power spectra of endplate current fluctuations

Muscle fibre type	τ noise (ms) Mean ± s.e. (range)	γ (pS) Mean ± s.e. (range)
Fast	4.78 ± 0.28 (3.8–5.89) $n = 7$	16.09 ± 0.58 (14.08–18.12) $n = 6$
Submaxillaris	6.87 ± 0.4 (5.9–8.85) $n = 9$	13.22 ± 0.73 (9.82–15.42) $n = 7$
Slow	14.8 ± 0.64 (12.3–17.68) $n = 8$	7.76 ± 0.57 (5.81–9.75) $n = 7$

n, Number of fibres; clamp potential, −90 mV; temperature 4–6 °C. τ, mean lifetime of channel; γ, mean conductance of single channel.

The time course of endplate currents induced by nerve stimulation was studied under membrane voltage clamp. Because both slow and submaxillaris muscle fibres have distributed innervation, it is necessary to discriminate synaptic events arising at the site of recording from those occurring elsewhere. One way of achieving this is to place the muscle in low Ca^{2+}–high Mg^{2+} solution and apply Ca^{2+} ionophoretically to a small region between the two microelectrodes used for clamping. In these conditions, nerve impulses continue to invade the nerve terminals but fail to release transmitter, unless an endplate lies close to the Ca^{2+} pipette, in which case release occurs but only from this junction (see ref. 12). To avoid further complications arising from time dispersal in the release of transmitter quanta[13], the efflux of Ca^{2+} from the pipette was adjusted so that nerve impulses released mainly single quanta. These unitary synaptic events, examples of which are shown in Fig. 2, were then used to measure the time course of transmitter action.

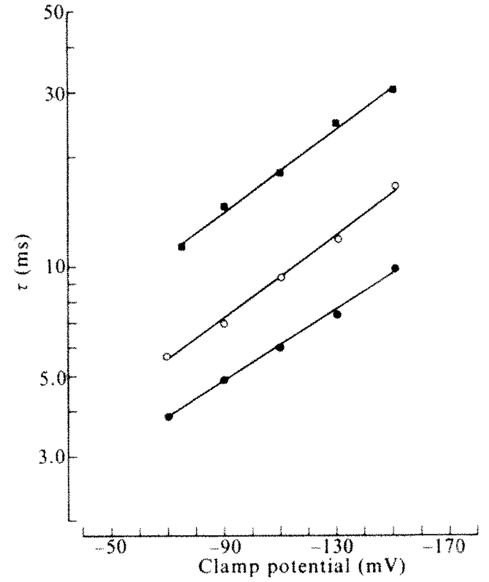


Fig. 3 Dependence of the time constant of decay of unitary endplate currents on membrane potential in a fast (●), a submaxillaris (○) and a slow (■) muscle fibre. Each point represents the mean of at least 10 unitary currents. An e-fold change in the time constant of decay was obtained when the muscle fibres were hyperpolarized by 84.5, 74.5 and 75.5 mV in the case of the fast, submaxillaris and slow muscle fibres, respectively. Normal Ringer at 6 °C.

The time constant of decay of unitary endplate currents, at a holding potential of −90 mV and at about 6 °C, was 5.0 ± 0.35 ms (mean ± s.e. of five fibres) in fast fibres and 14.67 ± 1.09 ms (six fibres) in slow fibres. Thus, the decay of transmitter action at slow muscle fibres is approximately three times slower than at fast fibres, a difference which resembles that found between fast and slow muscle fibres in the snake[14] and in the iliofibularis muscles[15]. In submaxillaris muscle fibres the decay constant of unitary endplate currents varied over a wide range (from 6.5 to 13.17 ms, mean 8.87 ± 0.56 ms, 13 fibres), even among fibres within the same muscle. The decay phase of the unitary endplate currents of the fast, submaxillaris and slow muscle fibres was prolonged as the membrane was hyperpolarized[16]. Moreover, as indicated by the slope of the voltage-decay time constant relation, the voltage sensitivity was approximately the same in the three types of fibres (Fig. 3). In contrast, the decay of spontaneous miniature endplate currents in slow fibres of the snake seems to be much less sensitive to voltage than in fast fibres[14]. We do not know if this difference is entirely due to the different species used (frog as compared with snake) or whether other factors are also involved.

To see if the different time courses of unitary endplate currents were due to differences in the lifetime of the channels opened by ACh in the various types of muscle fibres, we analysed the current fluctuations produced by ionophoretic application of ACh to the neuromuscular junction[7,8]. The mean lifetime (τ) of the junctional channels induced by ACh in submaxillaris muscle fibres was longer than in fast fibres, and the channel lifetime was longer still in slow fibres (Table 1). A close agreement was found between channel lifetime and the time constant of decay of unitary endplate currents in the three types of fibres, indicating that most of the ACh molecules, released from the motor nerve terminal at all these junctions, diffuse away from the ACh receptors or are destroyed by cholinesterase before they can exert repeated actions[17].

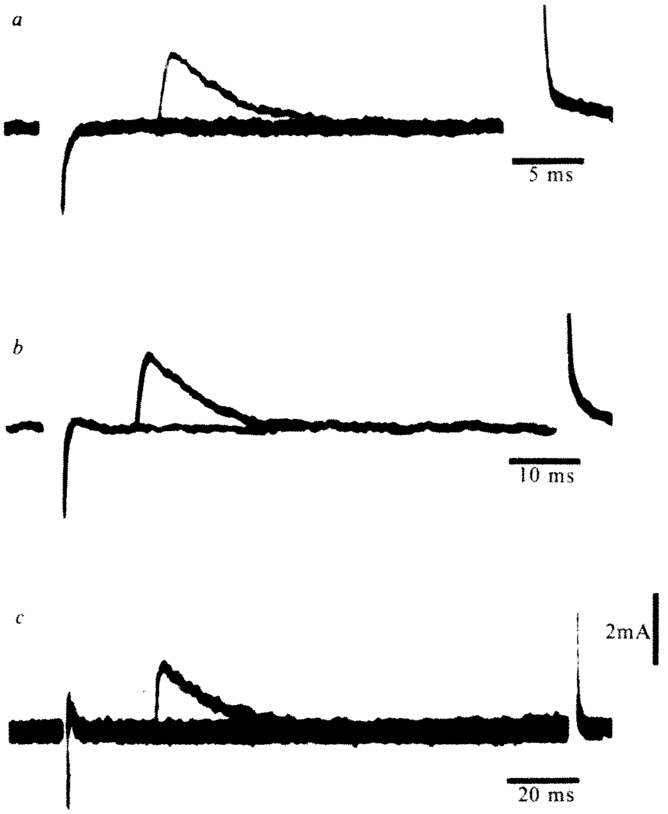


Fig. 2 Unitary (single quanta) endplate current responses to nerve stimulation, recorded from a fast (a), a submaxillaris (b) and a slow (c) muscle fibre. Muscles were in low Ca^{2+}–high Mg^{2+} solution at 6 °C. The artefact at the end of each trace indicates the start of the 800-ms pulse of Ca^{2+} used to restrict transmitter release to the endplate under study. Endplate currents were recorded under voltage clamp (−90 mV) through a 500-Hz low pass filter. The nerve was stimulated once every 1 or 2 s.

The equilibrium potential for ACh action was found to be near zero in the different types of fibres, and the single channel conductance (γ) was estimated from both the variance of the current noise and the zero frequency asymptote of the power spectra[7,8]. As shown in Table 1, junctional channels induced by ACh in fast fibres had larger conductance than the channels in submaxillaris fibres and these in turn had larger conductance than the channels in slow fibres. Here again, there seems to be a difference between frog and snake muscles, because fast and slow fibres in the snake appear to have the same single channel conductance[14].

Knowing the single channel conductance and the amplitude of the unitary endplate current, an estimate can be made of the number of channels opened by the release of one transmitter package[7]. The amplitudes of unitary endplate currents in fast, submaxillaris and slow fibres were, respectively (in nA), 1.98 ± 0.35 ($n = 6$), 1.76 ± 0.12 ($n = 13$) and 1.36 ± 0.16 ($n = 6$). This suggests that the mean number of channels opened by a single transmitter packet is ~1,370 in fast fibres, 1,480 in submaxillaris and 1,950 in the slow fibres. More experiments are needed to substantiate these preliminary estimates further, and to determine the cause of such a difference.

We conclude that the membrane channels activated by ACh at the three types of muscle fibres studied differ in the time they remain open and also in their conductance. Some properties of frog muscle fibres are known to be under neural control[18-22], and it is possible that the characteristics of the membrane channels induced by ACh on a given endplate are determined by the type of nerve which innervates it.

This work was supported by the Muscular Dystrophy Group of Great Britain. We thank Mr I. Parker for computing programming.

Received 12 December 1980; accepted 18 February 1981.

1. Taski, I. & Mizutani, K. *Jap. J. med. Sci. Biol.* **10,** 237–244 (1944).
2. Kuffler, S. W. & Vaughan Williams, E. M. *J. Physiol., Lond.* **121,** 289–317 (1953).
3. Smith, R. S. & Ovalle, W. K. *J. Anat.* **116,** 1–24 (1973).
4. Lännergren, J. *Nature.* **279,** 254–256 (1979).
5. Kordylewski, L. *Z. mikrosk-anat. Forsch.* **93,** 225–243 (1979).
6. Kordylewski, L. *Z. mikrosk-anat. Forsch.* **93,** 1038–1050 (1979).
7. Katz, B. & Miledi, R. *J. Physiol., Lond.* **224,** 665–700 (1972).
8. Anderson, C. R. & Stevens, C. F. *J. Physiol., Lond.* **235,** 655–692 (1973).
9. Ecker, A. &. Wiedersheim R. *Anatomie des Frosches* (Braunschweig, Vieweg, 1896).
10. Stefani, E. & Steinbach, A. B. *J. Physiol., Lond.* **203,** 383–401 (1969).
11. Burke, W. & Ginsborg, B. L. *J. Physiol., Lond.* **132,** 587–598 (1956).
12. Katz, B. & Miledi, R. *Proc. R. Soc.* **B161,** 496–503 (1965).
13. Katz, B. & Miledi, R. *J. Physiol., Lond.* **181,** 656–670 (1965).
14. Dionne, V. E. & Parsons, R. L. *Nature* **274,** 902–904 (1978).
15. Magazanik, L. G., Fedrov V. V. & Snetkov, V. A. *Prog. Brain Res.* **49,** 225–240 (1979).
16. Kordås, M. *J. Physiol., Lond.* **204,** 493–502 (1969).
17. Katz, B. & Miledi, R. *J. Physiol., Lond.* **231,** 549–574 (1973).
18. Miledi, R. *J. Physiol., Lond.* **151,** 1–23 (1960).
19. Elul, R., Miledi, R. & Stefani, E. *Acta physiol. latinoam.* **20,** 194–226 (1970).
20. Miledi, R., Stefani, E. & Steinbach, A. B. *J. Physiol., Lond.* **217,** 737–754 (1971).
21. Uchitel, O. D. *Acta physiol. latinoam.* **25,** 462–466 (1975).
22. Schmidt, H. in *Plasticity of Muscle* (ed. Pette, D.) (de Gruyter, Berlin, 1980).
23. Cordoba, F. *et al. J. Physiol., Lond.* **291,** 73P (1979).
24. Pearse, A. G. *Histochemistry, Theoretical and Applied* (Churchill Livingstone, London, 1972).

Relative proteolysis of the fibrinogen Bβ chain by thrombin and plasmin as a determinant of thrombosis

Hymie L. Nossel

Department of Medicine, College of Physicians & Surgeons of Columbia University, New York, New York 10032, USA

Thrombosis results, in part, from localized accumulation of fibrin, implying an imbalance between its rate of formation and dissolution. Astrup[1] postulated that patency of the vascular system depended on a dynamic equilibrium between constantly active coagulation and fibrinolytic systems. Reviews of this hypothesis have concluded that neither thrombin nor plasmin proteolysis makes a major contribution to fibrinogen turnover in normal individuals[2] and that the hypothesis of a dynamic equilibrium between clotting and lysis remains unproven[3]. Thus, the current view is that plasmin simply serves the function of a fibrinolytic enzyme digesting fibrin in the vascular system. Here, I promote the alternative view that the relative rates of proteolysis of the Bβ chain of fibrinogen by thrombin and plasmin determine the occurrence of thrombosis. This view is supported by recent data and can be used to make some readily testable predictions.

New biochemical information and technology now permit specific and quantitative analysis of the balance between thrombin and plasmin proteolysis of fibrinogen. Such analysis depends on measurement of the products of proteolysis which will briefly be reviewed. Thrombin cleaves fibrinogen to release sequentially fibrinopeptide A (FPA) and fibrinopeptide B (FPB)[4,5], implying two stages in fibrinogen–fibrin conversion with sequential activation of two sets of polymerization sites[6]. Removal of FPA exposes the A-binding domains in the NH_2-terminal region of fibrinogen which is now termed fibrin I. The A-domains of fibrin I bind to COOH-terminal domains termed 'a' in an adjacent molecule[7-10]. The removal of FPB produces fibrin II, in which a second set of binding sites is exposed—the B-domains which bind to the b-domains of an adjacent molecule. Laurent and Blombäck[11] observed that fibrin I polymers were less compact than those formed by fibrin II. Fibrin I was thought to form primarily end-to-end polymers and fibrin II both end-to-end and side-to-side polymers. Shen *et al.*[12] confirmed and extended these findings. They postulated that removal of FPA results in slow assembly of thin fibrin fibres and that rapid assembly to thick fibres depends on the removal of FPB. Hantigan and Hermans proposed that formation of protofibrils is a prerequisite for removal of FPB by thrombin, a step which is itself a prerequisite for rapid fibre formation and gelation[13]. In the early stages of plasmin proteolysis of fibrinogen, two-thirds of the COOH-terminal Aα chain and Bβ 1–42 from the NH_2-terminal Bβ chain are released to leave fragment X (refs 14–18). Hence, when fibrin I is the substrate, if the arginine (Bβ 14)–glycine (Bβ 15) bond is cleaved by thrombin, FPB is released and fibrin II formed, whereas if the arginine (Bβ 42)–alanine (Bβ 43) bond is cleaved by plasmin[17,19,20] the products are Bβ 1–42 and fragment X (Fig. 1). Fragment X

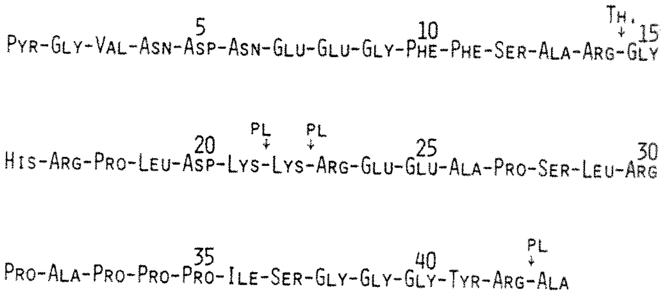

5
10
Th.
15
Pyr-Gly-Val-Asn-Asp-Asn-Glu-Glu-Gly-Phe-Phe-Ser-Ala-Arg-Gly
PL
PL
20
25
30
His-Arg-Pro-Leu-Asp-Lys-Lys-Arg-Glu-Glu-Ala-Pro-Ser-Leu-Arg
PL
35
40
Pro-Ala-Pro-Pro-Pro-Ile-Ser-Gly-Gly-Gly-Tyr-Arg-Ala

Fig. 1 Primary structure of the NH_2-terminal Bβ chain of fibrinogen. Bonds attacked by thrombin and plasmin are as indicated by Iwanaga *et al.*[19] and Wallen[20].

formed from fibrin I will form a clot but this is only 1% as rigid as that formed from intact fibrinogen[21], and α-polymer cross-links, a primary determinant of susceptibility to lysis by plasmin[22], do not form with fragment X[23,24]. Additionally, fibrin II is covalently cross-linked more rapidly than is fibrin I[25].

Some of these reactions can now be specifically quantified *in vivo*, as radioimmunoassays applicable *in vivo* have been developed for FPA reflecting fibrin I formation[26], thrombin-increasable FPB immunoreactivity (TIFPB) reflecting Bβ 1–42 and hence fragment X formation[27], and desarginine FPB reflecting fibrin II formation[28]. The mean normal plasma FPA level is 0.6 pmol ml^{-1} and elevated levels occur in acute venous and arterial thrombosis, pulmonary embolism, disseminated intravascular coagulation and other diseases[29–36]. It is evident that elevated FPA levels and hence fibrin I formation are not specific for thrombosis and that the effects of plasmin should also be considered. Peptide Bβ 1–42, reflecting the action of plasmin, can be identified in plasma as TIFPB[27]. In normal individuals the concentration of TIFPB is always higher than that of FPA, with a ratio of ~1.5–3 : 1. A reproducible sequence of changes in the concentration of FPA and TIFPB occurred in the blood of patients who received an intra-amniotic infusion of hypertonic saline to induce abortion[27]. First the FPA concentration increased and later the concentration of TIFPB rose, implying a sequence of fibrinogen proteolysis as shown in Fig. 2. If a single fibrinogen molecule is considered, in reaction (1), thrombin cleaves fibrinogen to yield free FPA and fibrin I monomer which polymerizes in reaction (2). Thrombin and plasmin then compete for the NH$_2$-terminal end of the Bβ chain (Fig. 1). If thrombin cleaves the Bβ chain before plasmin the products are FPB and fibrin II. If plasmin cleavage precedes that of thrombin the products are Bβ 1–42 and fragment X. Hence, plasmin cleavage of fibrin I serves as a major mechanism regulating the formation of fibrin II.

The facts that fibrin II forms more compact fibres than fibrin I and fibres with much greater tensile strength than fragment X and that elevated FPA levels are frequently found in conditions not associated with thrombosis prompt the hypothesis that the amount of fibrin II formed is important in thrombosis. The balance between the rates of fibrin II (reaction (3)) and fragment X (reaction (4)) formation would then determine the pathophysiological consequences of fibrin I generation. Predominant fibrin II formation results in thrombosis whereas predominant fragment X formation does not. The hypothesis explains and predicts phenomena, thus providing a means of testing its validity.

Phenomena predicted are that the ratio between peripheral blood FPA and TIFPB levels will demonstrate a greater association with thrombosis than FPA levels alone. Indices of fibrin II formation such as desarginine FPB levels may also be more closely associated with thrombosis than are indices of fibrin I formation. The thrombus itself should contain fibrin II. Although thrombosis does occur in the clinical syndrome of disseminated intravascular coagulation, there is often a marked dissociation between the degree of proteolysis of circulating fibrinogen and the extent of thrombosis found pathologically[37]. According to the above hypothesis, this discrepancy would result from a successful body defence mechanism for coping with a large amount of fibrin I formed—transformation to the less harmful fragment X. Hence, one would predict that much fragment X would be formed in disseminated intravascular coagulation and there are reports of the presence of this molecule in the plasma of patients with this disorder[38].

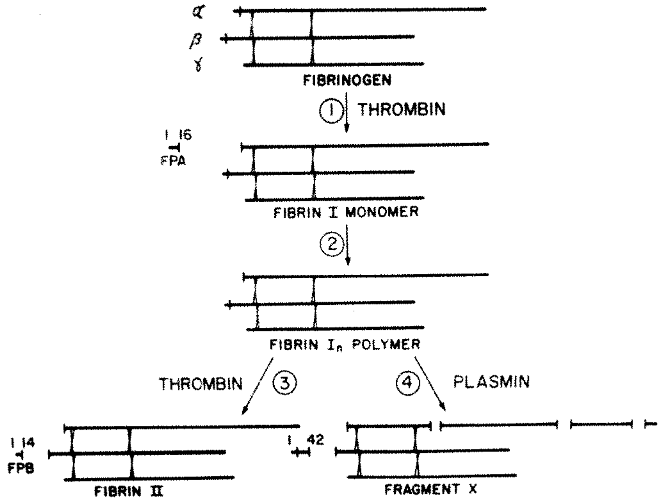


Fig. 2 Postulated scheme of fibrinogen proteolysis *in vivo*. Half fibrinogen molecules are depicted. Reaction (1), thrombin cleaves the Aα chain of fibrinogen to produce free FPA and fibrin I monomer. Reaction (2), polymerization occurs. Reaction (3), thrombin cleaves fibrin I to produce free FPB and fibrin II. Reaction (4), plasmin cleaves fibrin I to produce Bβ 1–42 and fragment X. Evidence that reaction (2) precedes reaction (3) is from Blombäck *et al.*[6]. Evidence that reaction (4) occurs preferentially to plasmin proteolysis of fibrinogen is from H.L.N. *et al.*[27]. Data have not been published on whether reaction (2) facilitates reaction (4).

The scheme shown in Fig. 2 implies that fragment X occurs as a polymer but the molecule seems to occur in monomeric form in plasma. Monomeric fragment X may be formed from depolymerization resulting from proteolysis or from plasmin proteolysis of fibrin I monomer. Ancrod and reptilase, which produce fibrin I[39,40] and not fibrin II, do not cause thrombosis and in fact are used in its therapy. Additional mechanisms for the lack of thrombosis with these enzymes are that Ancrod degrades the COOH-terminal portion of the Aα chain of fibrinogen and does not activate factor XIII and that neither enzyme aggregates platelets. Selectively increased thrombin concentration would favour reaction (3) and promote thrombosis, and congenital deficiency of antithrombin III, a condition expected to result in selective increase in thrombin concentration, is recognized as being associated with thrombosis[41,42]. Selectively decreased plasmin formation would decrease the rate of reaction (4), and conditions with decreased plasmin formation have also been reported to be associated with thrombosis[43–46]. Selectively increased plasmin activity due to decreased activity of α_2-antiplasmin may be associated with a bleeding tendency[47,48]. Finally, it is evident that although fibrin is an important component of thrombi, both platelets and the vessel wall constituents have important roles in thrombosis. Hence, the proposed hypothesis concerning fibrin formation and thrombosis deals with only a part of the process.

The work from my laboratory referred to here was supported by grants from the National Heart, Lung & Blood Institute (HL-15486, HL-21006 and HL-15596).

Received 19 December 1980; accepted 3 March 1981.

1. Astrup, T. *Blood* **11,** 781–806 (1956).
2. Hjort, P. F. & Hasselback, R. *Thromb. Diath. haemorrh.* **6,** 580–612 (1961).
3. McNicol, G. P. & Douglas, A. S. in *Human Blood Coagulation, Haemostasis and Thrombosis* 2nd edn (ed. Biggs, R.) 399 (Blackwell, Oxford, 1976).
4. Blombäck, B. & Vestermark, A. *Ark. Kemi* **12,** 173–182 (1958).
5. Bilezikian, S. B., Nossel, H. L., Butler, V. P. Jr & Canfield, R. E. *J. clin. Invest.* **56,** 438–445 (1975).
6. Blombäck, B., Hessel, B., Hogg, D. & Therkildsen, L. *Nature* **275,** 501–505 (1978).
7. Kudryk, B. J., Collen, D., Woods, K. R. & Blombäck, B. *J. biol. Chem.* **249,** 3322–3325 (1976).
8. York, L. L. & Blombäck, B. *Thromb. Res.* **8,** 607–618 (1976).
9. Kudryk, B., Blombäck, B. & Blombäck, M. *Thromb. Res.* **9,** 25–36 (1976).
10. Olexa, S. & Budzynski, A. *Proc. natn. Acad. Sci. U.S.A.* **77,** 1374–1378 (1980).
11. Laurent, T. C. & Blombäck, B. *Acta chem. scand.* **12,** 1975–1977 (1958).
12. Shen, L. L., Hermans, J., McDonagh, J. & McDonagh, R. P. *Am. J. Physiol.* **232,** H629–H633 (1977).
13. Hantigan, R. R. & Hermans, J. *J. biol. Chem.* **254,** 11272–12281 (1979).
14. Pizzo, S. V., Schwartz, M. L., Hill, R. L. & McKee, P. A. *J. biol. Chem.* **247,** 636–645 (1972).
15. Budzynski, A. Z., Marder, V. J. & Shainoff, J. R. *J. biol. Chem.* **249,** 2294–2302 (1974).
16. Mosesson, M. W. & Finlayson, J. S. in *Progress in Hemostasis and Thrombosis* Vol. 3 (ed. Spaet, T. H.) 61–107 (Grune & Stratton, New York, 1976).

17. Takagi, T. & Doolittle, R. F. *Biochemistry* **14,** 940–946 (1975).
18. Marder, V. J., Shulman, N. R. & Carroll, W. R. *J. biol. Chem.* **244,** 2111–2119 (1969).
19. Iwanaga, S., Wallen, P., Gröndahl, N., Henschen, A. & Blombäck, B. *Eur. J. Biochem.* **8,** 189–199 (1965).
20. Wallen, P. *Scand. J. Haemat. Suppl.* **13,** 3–14 (1971).
21. Shen, L. L., McDonagh, R. P., McDonagh, J. & Hermans, J. *J. biol. Chem.* **252,** 6184–6189 (1977).
22. Gaffney, P. & Brasher, M. *Biochim. biophys. Acta* **295,** 308–313 (1973).
23. Fretto, L. J., Ferguson, E. W., Steinman, H. M. & McKee, P. A. *J. biol. Chem.* **253,** 2184–2195 (1978).
24. Finlayson, J. S., Mosesson, M. W. & Bronzert, T. J. *J. biol. Chem.* **247,** 5220–5222 (1972).
25. Brosstad, F. *Thromb. Res.* **13,** 799–803 (1978).
26. Nossel, H. L. *et al. Proc. natn. Acad. Sci. U.S.A.* **68,** 2350–2353 (1971).
27. Nossel, H. L. *et al. J. clin. Invest.* **64,** 1371–1378 (1979).
28. Eckhardt, T. *et al. J. clin. Invest.* **67,** 809–816 (1981).
29. Nossel, H. L. *et al. J. clin. Invest.* **54,** 43–53 (1974).
30. Gerrits, W. B. J., Flier, O. T. N. & van der Meer, J. *Thromb. Res.* **5,** 197–212 (1974).
31. Theorell, H., Blombäck, M. & Kockum, C. *Thromb. Diath. haemorrh.* **36,** 593–604 (1976).
32. Nossel, H. L. *et al. J. clin. Invest.* **58,** 1136–1144 (1976).
33. Cronlund, M. *et al. J. clin. Invest.* **58,** 141–151 (1976).
34. Kockum, C. *Thromb. Res.* **8,** 225–236 (1976).
35. Budzynski, A. & Marder, V. J. *Thromb. Diath. haemorrh.* **34,** 709–717 (1975).
36. Yudelman, I. M., Nossel, H. L., Kaplan, K. L. & Hirsh, J. *Blood* **51,** 1189–1195 (1978).
37. Minna, J. D., Robboy, S. J. & Colman, R. W. in *Disseminated Intravascular Coagulation in Man* (ed. Thomas, C. C.) (C. C. Thomas, Springfield, 1974).
38. Merskey, C., Johnson, A. J., Harris, J. U., Wang, M. T. & Swain, S. *Br. J. Haemat.* **44,** 655–670 (1980).
39. Ewart, M. R., Hatton, M. W. C., Basford, J. M. & Dodgson, K. S. *Biochem. J.* **118,** 603–609 (1970).
40. Pizzo, S. V., Schwartz, M. L., Hill, R. L. & McKee, P. A. *J. clin. Invest.* **51,** 2841–2850 (1972).
41. Egeberg, O. *Thromb. Diath. haemorrh.* **13,** 516–630 (1965).
42. Marciniak, E., Farley, C. H. & DeSimone, P. A. *Blood* **43,** 219–231 (1974).
43. Astedt, B., Isacson, S., Nilsson, I. M. & Pandolfi, M. *Br. med. J.* **4,** 631–634 (1973).
44. Isacson, S. & Nilsson, I. M. *Acta chir. scand.* **138,** 313–319 (1972).
45. Nilsson, I. M., Krook, H., Sternby, N. H., Soderberg, E. & Soderstrom, N. *Acta med. scand.* **169,** 323–337 (1961).
46. Aoki, N., Moroi, M., Sakata, Y. & Yashida, N. *J. clin. Invest.* **61,** 1186–1195 (1978).
47. Koie, K., Tadashi, K., Ogata, K., Takamatsu, J. & Kobakura, M. *Lancet* **ii,** 1334–1336 (1979).
48. Aoki, N. *J. clin. Invest.* **63,** 877–884 (1979).

Epoxide hydrolase activity in the mitochondrial fraction of mouse liver

Sarjeet S. Gill & Bruce D. Hammock

Departments of Entomology and Environmental Toxicology,
University of California, Davis, California 95616, USA

The intense interest in the metabolic fate of epoxidized xenobiotics is due to several factors. For instance, epoxides are often intermediates in the lipophile to hydrophile conversions necessary for the excretion of olefinic and aromatic compounds by living systems[1], and are widely encountered in man's diet from both natural and man-made sources. Some of these epoxidized compounds may alkylate proteins and nucleic acids and thus include some of the most potent cytotoxins, mutagens and carcinogens known[2]. In mammals, epoxides may rearrange, deoxygenate to olefins, react with glutathione to form conjugates, or be hydrolysed by water to yield 1,2-diols with or without enzymatic catalysis[1,3,4]. The enzymes which catalyse the formation of diols are known as epoxide hydrolases (EC 3.3.2.3), and their subcellular distribution is the subject of this report. Early data showed that styrene oxide hydrolase activity was associated with the microsomal subcellular fraction[5]. Epoxide hydrolase activity was subsequently demonstrated on the nuclear[6], Golgi apparatus and plasma membranes[7], and in the cytosol of the cell[8,9], leaving the mitochondria as the last major cellular organelle assumed to be devoid of epoxide hydrolase activity. We now report strong evidence for the occurrence of substantial epoxide hydrolase activity in the mitochondria.

While studying the subcellular distribution of the microsomal epoxide hydrolase, Oesch *et al.*[5] reported substantial styrene oxide hydrolase activity in a crude mitochondrial fraction prepared in isotonic KCl. However, the same study showed that mitochondria prepared in isotonic sucrose had low styrene oxide hydrolase activity which was not significantly different from that of the microsomal marker enzyme glucose-6-phosphatase. Analogous subcellular distribution studies in this laboratory using an epoxidized terpene or fatty acid as substrate showed high epoxide hydrolase activity in the mitochondrial fraction[10,11]. However, in the case of these two substrates, the specific activity of the hydrolase was increased rather than decreased by repeated washing of the mitochondria. As no marker enzymes were used in these studies it was not possible to show definitively the presence of epoxide hydrolase activity in the mitochondria. Thus, the following experiments designed to yield 'pure mitochondria' were performed.

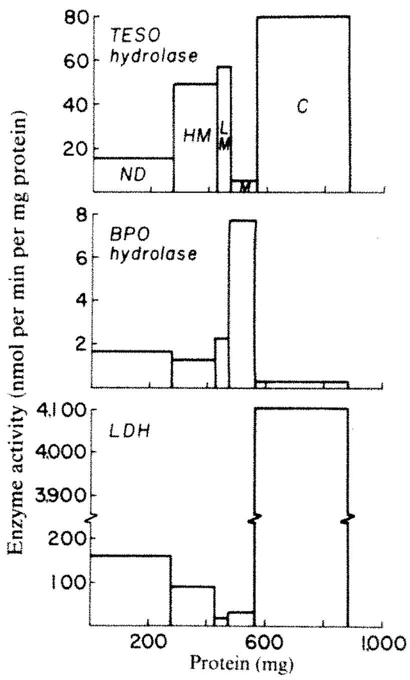


Fig. 1 Histograms of the subcellular distribution of epoxide hydrolase activity using TESO, *trans-β*-ethylstyrene oxide (upper), or BPO, benzo(*a*)pyrene 4,5-oxide (middle), as substrates and LDH, lactate dehydrogenase, with pyruvate as substrate (lower). Livers from 10-week old, male Swiss-Webster mice were rinsed in ice-cold 0.25 M sucrose and homogenized in 0.25 M sucrose/10 mM Tris-HCl, *p*H 7.4 buffer in a loose-fitting Potter-Elvejhem homogenizer to give a 10% w/v homogenate. The homogenate was centrifuged at 600*g* for 10 min, to give the nuclei and debris fraction (ND). The heavy mitochondria (HM) were pelleted at 6,500*g* for 10 min, the light mitochondria (LM) at 12,000*g* for 10 min, and the microsomes (M) at 100,000*g* for 60 min, leaving the cytosolic fraction (C) as the supernatant. Pelleted fractions were washed in the same conditions and resuspended in buffer (sodium phosphate, I = 0.2 M, *p*H 7.4 or Tris-HCl, 0.1 M, *p*H 9.0) for analysis. The ND, HM, LM, M and C fractions had 270, 160, 41, 87 and 320 mg of protein, respectively, from a preparation using three mouse livers. Separation of lysosomes and peroxisomes from mitochondria was achieved from livers, treated with Triton WR-1339 at 850 mg per kg 3.5 days before they were killed, by differential centrifugation as above followed by sucrose density gradient centrifugation[15]. The lysosomal and peroxisomal fractions had less than 2 and 28%, respectively, of the epoxide hydrolase activity in the crude mitochondrial fraction. The relatively high epoxide hydrolase activity in the peroxisomal fraction is probably due to mitochondrial contamination as the separation of these two fractions was not clear cut as determined by enzyme markers. Mitochondria were separated into inner membrane, outer membrane and the intermembrane and matrix fractions by swelling, shrinking and sonication, followed by sucrose density gradient centrifugation[16]. The values reported here are averages of two or three determinations from a preparation using three mouse livers. Of the total epoxide hydrolase activity present in the crude cell homogenate, 70 and 78% of the TESO and BPO hydrolysing activity is recovered in the various subcellular fractions.

Using the procedure described in Fig. 1 legend, the purity of the mitochondrial and submitochondrial fractions was monitored with several enzymes. Subcellular distribution of epoxide hydrolase activity was monitored with benzo(*a*)pyrene 4,5-oxide (BPO)[12] and *trans-β*-ethylstyrene oxide (TESO)[13] as

0028-0836/81/200167—02$01.00

substrates. Significant TESO hydrolase activity, ~25% of the total recovered in the subcellular fractions, was observed in the mitochondrial fraction (Fig. 1), although the highest level was found in the cytosol. The subcellular distribution of epoxide hydrolase activity is very similar when *trans*-stilbene oxide (TSO) is used as substrate. Both TESO and TSO are negligibly hydrated by the microsomal epoxide hydrolase and the immunologically similar nuclear epoxide hydrolase. In contrast, the mitochondrial fraction poorly hydrates BPO, a substrate widely used to monitor epoxide hydrolase activity. Further, the low specific activity of BPO hydrolase in the mitochondrial fraction indicates minimal contamination of this fraction by nuclear and/or microsomal membranes as BPO hydrolase is a good marker enzyme for the microsomal and nuclear membranes[1,6]. Repeated washing of the mitochondrial fractions leads to a substantial decrease in BPO hydrolase specific activity and most of this activity is associated with the outer mitochondrial membrane. As styrene oxide and BPO hydrolase activity in the microsomes are immunologically identical[14], our data support the hypothesis of Oesch[1] that styrene oxide hydrolase activity in the mitochondrial fraction is probably due to microsomes adhering to the outer mitochondrial membrane. In contrast, the TESO hydrolase specific activity in the mitochondrial fractions increases following repeated washings.

Using lactate dehydrogenase, a widely accepted marker for the cytosol, we demonstrated that epoxide hydrolase activity in the mitochondrial and nuclear fractions does not result from cytosolic contamination (Fig. 1). Results using additional marker enzymes including cytochrome oxidase[15], NADPH cytochrome *c* reductase[16], cytochrome P_{450} (ref. 17), citrate synthase[18] and adenylate kinase[19] also support the relative purity of the subcellular fractions. Use of citrate synthase and adenylate kinase also showed that the cytosolic epoxide hydrolase activity does not arise from mitochondrial lysis during homogenization, as only 0.6 and 2% of the mitochondrial activity of these enzymes, respectively, was found in the cytosol.

Subsequent experiments using livers from mice treated with Triton WR-1339 (ref. 15) indicated that lysosomes and peroxisomes, monitored by acid phosphatase[19] and urate oxidase[20], respectively, have minimal TESO hydrolase activity. As such, epoxide hydrolase activity in the mitochondrial fraction does not arise from contamination by these organelles. When mitochondria are disrupted, a preponderance of the epoxide hydrolase activity is associated with the matrix and intermembrane space fraction with TESO as substrate and less activity is associated with the inner and outer mitochondrial membranes. Most of the activity associated with either of these membranes can be removed with a single wash, indicating contamination by the matrix and intermembrane space fraction. The mitochondrial fraction is able to hydrate a wide variety of epoxidized compounds including epoxidized fatty acids and glycidal ethers, with a spectrum of activity similar to that of the cytosolic epoxide hydrolase, but distinct from that of the microsomal enzyme.

As the nuclei and mitochondria are excluded from the activation/deactivation systems used in many short-term mutagenicity assays[21], these systems may not provide a valid picture of the ability of cells to degrade potentially mutagenic epoxides. However, this effect may be more quantitative than qualitative, because the mitochondrial epoxide hydrolase demonstrates a spectrum of activity similar to that of the cytosolic epoxide hydrolase.

The present study and published literature[1,6,7,22] illustrate the apparently ubiquitous nature of epoxide hydrolases in cellular components. Although the information needed to understand fully the biological role of epoxide hydrolases is lacking, their ability to metabolize some xenobiotics is clear. The association of epoxide hydrolases with the mitochondria and nuclei of the cell may serve to protect the structural integrity of these organelles as well as their nucleic acid content. It was recently shown[23,24] that mitochondrial DNA is more easily alkylated by polyaromatic hydrocarbons and their diol-epoxides than nuclear DNA. Such results are not altogether surprising because of the nature of mitochondrial DNA and the oxidases occurring in the mitochondria[25,26] which can activate the polyaromatic hydrocarbons to reactive electrophiles. Although the mitochondrial epoxide hydrolase activity reported here does not appreciably metabolize BPO, these epoxide hydrolases can potentially protect the mitochondria from those reactive epoxides which are substrates. Therefore, our study certainly illustrates the need for increased research on these important enzymes using a variety of potential substrates in all subcellular components.

This work was supported by grants from the National Institute of Environmental Health Sciences. ^{3}H-BPO was supplied by F. Oesch.

Received 20 October 1980; accepted 23 February 1981.

1. Oesch, F. *Xenobiotica* **3,** 305–340 (1973).
2. Miller, J. A. & Miller, E. C. in *Chemical Carcinogenesis* Pt A (eds Ts'o, P. O. P. & DiPaolo, V. A.) 61–85 (Dekker, New York, 1974).
3. Ivie, G. W. *Science* **191,** 959–961 (1976).
4. Jakoby, W. B. *Adv. Enzym.* **46,** 383–414 (1978).
5. Oesch, F., Jerina, D. M. & Daly, J. *Biochim. biophys. Acta* **227,** 685–691 (1971).
6. Mukhtar, H., Elmamlouk, T. H. & Bend, J. R. *Archs Biochem. Biophys.* **192,** 10–21 (1979).
7. Stasiecki, P., Oesch, F., Bruder, G., Jarasch, E.-D. & Franke, W. W. *Eur. J. Cell Biol.* **21,** 79–92 (1980).
8. Gill, S. S., Hammock, B. D. & Casida, J. E. *J. agric. Fd Chem.* **22,** 386–395 (1974).
9. Hammock, B. D., Gill, S. S., Stamoudis, V. & Gilbert, L. I. *Comp. Biochem. Physiol.* **53B,** 263–265 (1976).
10. Gill, S. S. & Hammock, B. D. *Biochem. Pharmac.* **29,** 389–395 (1980).
11. Gill, S. S. & Hammock, B. D. *Biochem. biophys. Res. Commun.* **89,** 965–971 (1979).
12. Jerina, D. M., Dansette, P. M., Lu, A. Y. H. & Levin, W. *Molec. Pharmac.* **13,** 342–351 (1977).
13. Mullin, C. A. & Hammock, B. D. *Analyt. Biochem.* **106,** 476–485 (1980).
14. Oesch, F. & Bentley, P. *Nature* **259,** 53–55 (1976).
15. Leighton, F. *et al. J. Cell Biol.* **37**, 482–513 (1968).
16. Sottocassa, G. L., Kuylenstierna, B., Ernster, L. & Bergstrand, A. *J. Cell Biol.* **32,** 415–438 (1967).
17. Omura, T. & Sato, R. *J. biol. Chem.* **239,** 2370–2378 (1964).
18. Shepherd, D. & Garland, P. B. *Meth. Enzym.* **13,** 11–16 (1969).
19. Bergmeyer, H. U. *Methods of Enzymatic Analysis* Vols 1, 2 (Academic, New York, 1974).
20. Kalckar, H. M. *J. biol. Chem.* **167,** 461–475 (1947).
21. Ames, B. N., McCann, J. & Yamasaki, E. *Mutat. Res.* **31,** 347–364 (1975).
22. Hammock, B. D., Gill, S. S., Mumby, S. M. & Ota, K. in *Molecular Basis of Environmental Toxicity* (ed. Bhatnagar, R. S.) 229–272 (Ann Arbor Science, Michigan, 1980).
23. Allen, J. A. & Coombs, M. M. *Nature* **287,** 244–245 (1980).
24. Backer, J. M. & Weinstein, I. B. *Science* **209,** 297–299 (1980).
25. Sato, R., Atsuta, Y., Imai, Y., Taniguchi, S. & Okuda, R. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5477–5481 (1977).
26. Raw, I. *Biochem. biophys. Res. Commun.* **81,** 1294–1297 (1978).

Errata

In the letter 'Human immunoglobulin variable region genes—DNA sequences of two V_κ genes and a pseudogene' by D. L. Bentley and T. H. Rabbitts, *Nature* **288,** 730–733 (1980), Fig. 3 shows the DNA sequence of a human V_κ pseudogene (not V_λ).

In the article 'Sequence and organization of the human mitochondrial genome' by S. Anderson *et al.*, *Nature* **290,** 457–465, line 56 on page 462 should read '. . . of the anticodon of mammalian mt RNAs is *not* known, it is . . .'

In the article 'Identification of D segments of immunoglobulin heavy-chain genes and their rearrangement in T lymphocytes' by Y. Kurosawa *et al.*, *Nature* **290,** 565–570, the arrows indicating recombination sequences in the text figure on page 570 were positioned incorrectly. The correct version is shown below:

```
                     B6       Q52
                      ↓        ↓
DQ52      GCACCACAGTGCAACTGGGACCACTGTGACGC
JH2       GATAGTGTGACTACTTTGACTACTGGGGCCAA
                      ↑      ↑
                     B6     Q52
```

Corrigendum

In the letter 'A nonrandom component in cosmic rays of energy $\geqslant 10^{14}$ eV' by C. L. Bhat *et al.*, *Nature* **288,** 147–149 (1980), the RA values shown in Figs 2 and 3 have been found to be incorrect. The correct values are those obtained by subtracting the values given from 24 hours. This excludes the pulsar PSR 0525+21 as a possible origin. The results are consistent with a point-source origin in the RA range 20±3 hours.

MATTERS ARISING

Fortuitous oxidations by methane-utilizing bacteria

A MAJOR thesis of Higgins *et al.*[1] is that the wide substrate specificity shown by methane monooxygenase (MMO) has developed because of its survival value to the species by enabling the organisms to use ancillary carbon or energy sources.

In support of their thesis, the authors state that "such extraordinary lack of substrate specificity would be extremely unusual if it were entirely fortuitous". There are, however, many examples of oxygenases and other enzymes which have a wide substrate range but which are only metabolically functional for a single substrate. For example, ribulose bisphosphate carboxylase in all autotrophs has both carboxylation and oxygenation activities, the latter apparently leading to wasteful metabolism which has not been eliminated over a long period of biochemical evolution[2]. Nitrogenase, an enzyme which reduces many triply bonded substrates, also produces dihydrogen as an intrinsic part of the reduction process, clearly a wasteful mechanism[3]. If there were some survival value to the species there would be a strong selective pressure on the organism to eliminate these 'side' reactions or to retain, evolve or acquire those enzymes that were capable of assimilating the product into biomass. Although it is possible that acetate could arise as a result of β-oxidation of 1-phenylheptane and ω-oxidation of higher alkanes and may therefore subsequently yield reducing equivalents, there is no direct evidence that any of the biotransformation products given in ref. 1, Table 1 are incorporated into biomass. Indeed, the biotransformation of most of these substrates results in a net loss of available energy to the cell because they require NADH for oxygenation to the product.

It is possible, for example, that the oxidation of toluene to benzoic acid may result in a net gain in reducing equivalents due to the activity of nonspecific alcohol dehydrogenases; however, no NADH is generated. As MMO requires NAD(P)H *in vitro*[4–6] (and probably *in vivo* also), any oxidation that does not give rise to NAD(P)H may well be detrimental to the cell's overall energy balance. Furthermore, it has been calculated that the NAD(P)H-requiring MMO organisms are probably limited for NAD(P)H supply during growth on methane[7]. The fact that some of the products are certainly more toxic to the cell than the substrate also argues against the retention of a wide-specificity MMO. We therefore contend that these biotransformations, which are supposed to aid the survival of the organisms, probably have exactly the opposite effect.

On closer examination of the paper quoted by the authors on their own work using whole cells, one observes that they were using suspensions containing 70–80 mg dry weight of organisms per 20 ml to effect the biotransformations of these compounds over a 15–24-h period. The high cell concentration necessary to detect oxidation is clearly unrepresentative of the situation in the biosphere.

Therefore, we feel that the arguments for the wide substrate specificity of MMO having "substantial metabolic significance" are invalid. We re-affirm our belief that all these oxidations are purely fortuitous[8] and that there is no direct evidence for the oxidation of any substrate by MMO, other than methane, conferring a special advantage to the cell, whether it be as a supplementary carbon source or a provider of useful energy. We have always subscribed to the view that methanotrophs have a significant role in the oxidation of organic compounds and ammonia in the environment, but purely by a fortuitous mechanism and not one that has evolved or been retained to confer a substantial metabolic advantage to the cell.

Some methanotrophs, faced with a shortage of methane in the environment, may well oxidize any available ethane to acetate, but significant amounts of such acetate could be incorporated only by the type II methanotrophs (and not type I methanotrophs as the latter have an incomplete tricarboxylic acid cycle). Acetate could only serve as a source of carbon and not ATP because these organisms are unable to grow on acetate alone. Even so, it would still be necessary to generate NADH for ethane oxidation from the oxidation of methane to CO_2. But why is it necessary to do this for survival? All the methane oxidizers (where explored) seem to be able to form resting cells (cysts or exospores) which can survive in a hypobiotic state for years. Surely the elaboration of these protective mechanisms is far more relevant to survival than a catholic MMO enzyme able to scavenge trace amounts of carbon from the environment for the doubtful pleasure of making acetate?

D. I. STIRLING
H. DALTON
Department of Biological Sciences,
University of Warwick,
Coventry CV4 7AL, UK

1. Higgins, I. J., Best, D. J. & Hammond, R. C. *Nature* **286**, 561–564 (1980).
2. Ellis, R. J. *Trends biochem. Sci.* **4**, 241–244 (1980).
3. Robson, R. L. & Postgate, J. R. *A. Rev. Microbiol.* **34**, 183–207 (1980).
4. Colby, J., Dalton, H. & Whittenbury, R. *Biochem. J.* **151**, 459–462 (1975).
5. Colby, J. & Dalton, H. *Biochem. J.* **157**, 495–497 (1976).
6. Stirling, D. I. & Dalton, H. *Eur. J. Biochem.* **96**, 205–212 (1979).
7. Anthony, C. *J. gen. Microbiol.* **104**, 91–104 (1978).
8. Stirling, D. I. & Dalton, H. *FEMS Microbiol. Lett.* **5**, 315–318 (1979).

HIGGINS *ET AL.* REPLY—We do not accept the argument that there is nothing unusual about the extreme lack of specificity of some methane monooxygenases (MMO). Most oxygenases have a relatively narrow substrate range, some of them only acting on one or two very similar compounds. Even in the case of enzymes closely related to MMO, the microbial alkane hydroxylases, each one is restricted to quite a narrow band of chain lengths. The only oxygenase of which we are aware that approaches MMO in its 'catholic taste' for substrates is the mammalian liver P450 LM fraction[1]. Here it is very probable that there has indeed been selection for broad specificity as this enzyme is essential to initiating processes for the removal of a wide range of toxic metabolites and xenobiotics from the body. In addition, in some methanotrophs the MMO seems to have a more restricted substrate range[2], showing that extreme lack of specificity is not a mechanistic dictum.

We accept that the lack of specificity of ribulose bisphosphate carboxylase and nitrogenase leads to wasteful metabolism in some circumstances and various species have developed mechanisms to minimize this in the former case. Nevertheless, compared with MMO these are highly specific enzymes. Indeed, ribulose bisphosphate carboxylase is specific for O_2 or CO_2.

In reply to the specific points raised by Stirling and Dalton:

(1) It is pointed out that many of the reactions catalysed are actually detrimental to the organisms because of reducing power consumption or the generation of toxic products. This point was made in our review. If this is the case, lack of development of more specific enzymes would seem to require some explanation. In fact, of course, this is something of a 'red herring' because most compounds of that type are not present in the environment in substantial amounts. Some have only recently been introduced as a result of the activities of the organic chemist. Therefore, we agree that many of these actual oxidations are probably purely fortuitous, although the ability to carry out some of these oxidations, we suggest, has been retained as a desirable characteristic. We did not imply that most of the reactions in our Table 1 lead to utilizable carbon or energy generation. The purpose of the table was simply to give an idea of the types of reactions catalysed. Such reactions may be of industrial value and may be involved in

0028–0836/81/200169—02$01.00

oxidation of a range of compounds in the environment.

The only substrates for MMO of significance *vis à vis* any selective value for broad-specificity enzymes are those which on subsequent metabolism yield acetate. There are two reasons for this. First, these compounds (long-chain alkanes, long-chain alkenes, phenylalkanes, carotenoids) are present in very substantial quantities in the environment; they are major components of biomass. Second, acetate once formed can act as a carbon source for methanotrophs and in the case of type II species which possess an intact tricarboxylic acid cycle, acetate can be oxidized to CO_2 yielding reducing power and ATP. It is well authenticated that acetate is extensively metabolized by type II species and leads to appreciable increases in growth yields when included in the medium[3]. Phenylalkanes are partially degraded by *Methylosinus trichosporium* to short-chain alkanoic acids and alkanes are oxidized to CO_2. The most likely mechanism for this is β-oxidation, the product of which is acetate. Further evidence for β-oxidation comes from the finding that C-odd phenylalkanes yield only C-odd phenylalkanoic acids and C-even phenylalkanes yield mainly C-even phenylalkanoic acids[4]. The example of 1-phenylheptane metabolism was given in our Table 1.

(2) There is accumulating evidence that methanotrophs possess other broad-specificity enzymes including secondary alcohol and aldehyde dehydrogenases[5,6] that facilitate the further oxidation of products of MMO activity. Some of these enzymes, such as the secondary alcohol dehydrogenase discovered by the Exxon group, have no known function in C-1 metabolism.

(3) The environmental significance of reactions observed in our experiments using high cell densities is questioned. Of course, it is always difficult to extrapolate from laboratory experiments to the environment. It is not necessary to use such high cell densities to detect products, but clearly at very low cell densities product detection is difficult for technical reasons. We have used high cell densities to obtain in a short time sufficient material from small-scale incubations for both detection and chemical identification. In addition, because some of these substances are toxic at high concentrations, appropriate ratios of organisms to test compounds are necessary.

(4) We agree that the formation of exospores or cysts has great survival value. Would it not be advantageous, however, to be able to survive in a vegetative state using ancillary carbon and energy sources while your competitors are encysting?

(5) The fact that methanotrophs do not grow on acetate alone is not proof that acetate can act only as a carbon source. Acetate substantially increases yields of type II organisms on methane, suggesting that NAD(P)H and/or ATP are generated from this compound. It seems likely that type II methanotrophs have retained an intact tricarboxylic acid cycle to facilitate acetate oxidation. The reasons for the inability to grow on acetate may well be found in the mechanisms of regulation of intermediary metabolism.

There are many examples of microbial co-oxidations in which ATP or NADH are generated from substances which are not growth substrates. Methanotrophs will not grow on formate but generate NADH (and possibly ATP) from this intermediate.

We suggest that more experimental work should be done to test our hypothesis that lack of specificity of MMO leads to useful supplementary metabolism and this is in progress in our laboratory.

Note added in proof: Dalton and colleagues have previously subscribed to the view that in the presence of methane small amounts of co-oxidizable compounds may well contribute significantly to the growth of methanotrophs and not just be merely oxidized without benefit to the organism[7,8].

I. J. HIGGINS
D. J. BEST
R. C. HAMMOND
Biological Laboratory,
The University,
Canterbury, Kent CT2 7NJ, UK

1. White, R. E. & Coon, M. J. A. *Rev. Biochem.* **49,** 315–356 (1980).
2. Stirling, D. I. & Dalton, H. *FEMS Microbiol. Lett.* **5,** 315–318 (1979).
3. Wadzinski, A. M. & Ribbons, D. W. *J. Bact.* **123,** 380–381 (1975).
4. Higgins, I. J., Scott, D. & Best, D. *Proc. 3rd. int. Symp. Microbial Growth on C-1 Compounds* (ed. Dalton, H.) (Heyden, London, in the press).
5. Hou, C. T. *et al. Appl. envir. Microbiol.* **38,** 135–142 (1979).
6. Patel, R. N. *et al. Archs Biochem. Biophys.* **203,** 654–662 (1980).
7. Whittenbury, R. *et al.* in *Proc. 1st. int. Symp. Microbial Growth on C-1 Compounds* (ed. Terui, G.) 1–9 (Society of Fermentation Technology, Tokyo, 1975).
8. Whittenbury, R. *et al.* in *Symp. Microbial Production and Utilization of Gases* (eds Schlegel, H. G., Gottschalk, G. & Pfenning, N.) 281–292 (Goltze, Göttingen, 1976).

Cycloheximide inhibits DNA synthesis in SV40-transformed cells

IDE *ET AL.*[1] reported that DNA synthesis in an SV40-transformed human cell line (SV80) is not inhibited by the protein synthesis inhibitor cycloheximide. This is an apparent exception to the repeated observation that inhibitors of protein synthesis are effective inhibitors of DNA and RNA (mostly ribosomal) synthesis[2,3]. Using various other cell types, both neoplastic and normal, we have found[4,5] that the inhibitory effect of these drugs is also exerted on several reactions in both the salvage and *de novo* pathways of nucleotide biosynthesis, such as the formation of CTP from uridine and aspartate and the formation of GTP from hypoxanthine, adenosine, deoxyadenosine and formate. CTP and GTP are not only required for RNA synthesis but are also intermediates in the formation of the deoxynucleoside triphosphate precursors, dCTP, dTTP and dGTP. Various lines of evidence emphasize the importance of maintaining normal pool sizes of these precursors for RNA and DNA synthesis[6,7], particularly for the latter, as the level of the deoxynucleoside triphosphates is very small in eukaryotic cells.

The SV40-transformed system could be pertinent to the question of whether DNA synthesis inhibition depends solely on the synthesis of some unknown protein(s) of short half life, in some way necessary for the polymerization reaction, or additionally on the levels of newly synthesized precursors. Thus, we had hoped to use the transformed cells to determine whether, in conditions in which DNA synthesis is not affected, the reactions leading to the biosynthesis of the precursors are similarly unaffected. Unfortunately, we are unable to confirm the observations of Ide *et al.* In our study, after addition of cycloheximide or puromycin in both the human SV80 line that Ide *et al.* used, and in an SV40-transformed mouse line (BALB/c SVA31), DNA synthesis was inhibited to the same extent as in the untransformed cell (BALB/c A31). In addition, all the effects on the nucleotide pool constituents that we observed with other cell types were seen with the SV40-transformed lines.

There is no apparent explanation for the differences between our results and those of Ide *et al.*, who suggest that the persistence in DNA synthesis in transformed cells is associated with the T-antigen content of SV40-transformed lines. T-antigen assays of our stocks showed that the BALB/c SVA31 and the SV80 lines are strongly positive whereas the untransformed BALB/c A31 is negative. We conclude that SV40-transformed cells are not exceptions to the statement[8] that "inhibition of protein synthesis by any mechanism results in almost immediate interruption of DNA replication."

ELLIOT VOLKIN
PETER BENSON
Biology Division,
Oak Ridge National Laboratory,
and the University of
Tennessee–Oak Ridge Graduate School of Biomedical Sciences,
Oak Ridge, Tennessee 37830, USA

1. Ide, T., Ohta, T., Shinohara, Y. & Ishibashi, S. *Nature* **278,** 264–265 (1979).
2. Sisler, H. D. & Siegel, M. R. in *Antibiotics I: Mechanism of Action* (eds Gottlieb, D. & Shaw, P. D.) 283–307 (Springer, New York, 1967).
3. Muramatsu, M., Shimada, N. & Higashinakagawa, T. *J. molec. Biol.* **53,** 91–106 (1970).
4. Pike, L. M., Khym, J. X., Jones, M. H., Lee, W. H. & Volkin, E. *J. biol. Chem.* **255,** 3340–3347 (1980).
5. Volkin, E., Boling, M. E., Jones, M. H., Lee, W. H. & Pike, L. M. *J. biol. Chem.* **255,** 9105–9109 (1980).
6. Grummt, I. & Grummt, F. *Cell* **7,** 447–453 (1976).
7. Skoog, K. L., Bjursell, K. G. & Nordenskjold, B. A. *Adv. Enzyme Regul.* **12,** 343–354 (1974).
8. Sheinen, R., Humbert, J. & Pearlman, R. E. *A. Rev. Biochem.* **47,** 277–316 (1978).

0028–0836/81/200170—02$01.00

Beaming of Jupiter's decametric radiation

POQUÉRUSSE AND LECACHEUX[1] reported the first direct measurement of the beaming of Jupiter's decametric radiation based on simultaneous observations of the jovian radio emission made with the Soviet spacecraft Mars 7, and at the ground level at Nancay, France on two different days. On one of them, 15 May 1974, activity was recorded at the spacecraft while none was observed at Nancay or at the University of Colorado site, and it was concluded that this effect was evidence of the narrow beaming of the jovian decametric radiation. We made complementary observations from the Maipú Radio Astronomy Observatory in Chile and from the University of Florida Radio Observatory: strong activity from Jupiter was recorded at the same time at a lower frequency and at a close frequency some 40 min earlier.

Table 1 shows the time and frequencies where the activity was recorded, including the spacecraft, Florida and Maipú. In Florida the emission was detected at 20.0 and 22.2 MHz, while in Maipú, Chile at 13.1, 16.7, 18.0, 22.2 and 27.6 MHz. jovian activity was recorded at 22.2, 18.0 and 16.7 MHz and at the same time the spacecraft detected activity at 30 MHz while that at 27.6 MHz was recorded some 40 min earlier.

The antennas used were linearly polarized Yagis except for those operating at 13.1, 16.7 and 22.2 MHz in Chile, which were crossed polarimeter Yagis. The receiving, recording and calibrating system at both observatories were identical, each channel consisting of a modified commercial communication receiver tuned to the corresponding centre frequency with a fixed bandwidth of 4 kHz, a chart recorder with a time constant close to 1 s was used, although a supplementary recording was available for better resolution of individual bursts. Temperature-limited noise diodes were used as standard calibrator. Occasionally, no calibration was made during the observation—the galactic background noise serving as an auxiliary calibrator.

The flux density for each channel was determined following the standard procedure of the Florida–Chile group[2,3].

The flux densities of mean and peak pulses calculated for the observed activity at Maipú and Florida during 15 May 1974 are also included in Table 1. The emission recorded was mostly of the S-bursts type (millisecond pulses) and right-hand circularly polarized (16.7 and 22.2 MHz, Chile). The minimum detectable flux density (threshold) at Maipú and Florida was of the order of 2.5×10^{-22} W m^{-2} Hz^{-1} at 22.2 MHz. The threshold of the University of Colorado spectrograph (7–41 MHz) was reported to be 5×10^{-22} m^{-2} Hz^{-1} (ref. 1), that is well below the flux density of any mean pulse recorded in Chile or Florida (see table 1). The lack of correlation between the data obtained in Maipú and Florida suggests a strong ionospheric disturbance along the line of sight of Jupiter as seen from Florida. The same disturbance also seems to have affected the observations at Colorado, where Jupiter was being observed at a large zenith angle. It might even be possible to consider some scattering produced by field-aligned ionospheric irregularities.

Table 1 Summary of observations, 15 May 1974

Site	Frequency	Activity time UT		Flux density (10^{-22} Wm^{-2} Hz^{-1})		Remarks
Spacecraft	30	10.35	11.06	—	80	Ref. 1
Nancay	25–38	—	—	—	—	Ref. 1
Colorado	7–45	—	—	—	—	Observing time: 09.00–12.00
Florida	20.0	11.26	12.00	12.7	48.2	Jupiter transit: 13.03
	22.2	11.19	12.13	21.5	83.0	
Maipú	13.1	10.20	10.40	52.2	205.9	Jupiter transit: 12.11; sunrise: 11.26
	16.7	09.40	11.27	76.3	172.1	Right-hand polarized
	18.0	09.40	11.33	23.2	52.0	
	22.2	09.42	11.25	19.9	45.6	Right-hand polarized
	27.6	09.45	09.55	50.0	202.1	Antenna breakdown at 09.55

Observing frequencies, beginning and ending times of the activity, and mean and peak flux density measurements are given for each site where activity was recorded. At the highest frequency in Maipú (27.6 MHz), the antenna had a major breakdown at 09.55 so that the jovian activity was only recorded up to ~40 min before the spacecraft started recording the emission from Jupiter.

We can thus add to Poquérusse and Lecacheux's result that Jupiter's strong beaming may hold only for a relatively narrow frequency bandwidth. As might be expected, the radiation beam appears wider when all frequencies are considered rather than a single frequency.

FRANCISCO REYES*
Department of Physics
and Astronomy,
University of Florida,
Gainesville, Florida 32611, USA

JORGE MAY
Observatorio Radioastronómico,
Universidad de Chile,
Maipú, Chile

* Permanent address: Departmento de Astronomiá, Universidad de Chile, Maipú, Chile.

1. Poquérusse, M. & Lecacheux, A. *Nature* **275**, 111–113 (1978).
2. Carr, T. D. *et al. Astrophys. J.* **140**, 778–795 (1964).
3. May, J. & Carr, T. D. *Publ. Dept. Astr. Univ. Chile* No. 6, 87–92 (1969).

Corticosteroid-binding globulin and corticosterone interaction with progesterone receptors

IN their article[1] Al-Khouri and Greenstein claim to have shown "differences between progesterone receptors and CBG [corticosteroid-binding globulin] distribution in brain and uterus". We would like to point out that, in their studies, they have failed to take into consideration binding of ^{3}H-progesterone to other types of receptor sites, in particular glucocorticoid sites, and that for this reason they cannot conclude that "progesterone, the natural hormone, displaces the tritiated steroid from the receptor in brain and uterine cytosols more potently than does R5020".

In the conditions used by Al-Khouri and Greenstein (non-ovariectomized animals, not primed with oestrogens), progestin-binding sites (PR) are occupied by endogenous progesterone which cannot be exchanged with ^{3}H-progesterone within 2 h at 0 °C without loss of sites by degradation. Furthermore, even in the perfused brain and in optimal experimental conditions, the number of PR is low compared with the concentration of glucocorticoid binding sites (GR) and CBG. In the non-perfused uterus the relative quantity of PR compared with GR and CBG must be even smaller and ^{3}H-progesterone binding corresponds to the total binding of PR, GR and CBG and not to PR uniquely. Thus, in the displacement studies, cold R5020 displaces ^{3}H-progesterone primarily from PR and to a lesser extent from GR (note that competition of progesterone for GR is underestimated in screening assays because of trapping of progesterone by CBG) and cold corticosterone displaces ^{3}H-progesterone from GR and CBG. This explains why corticosterone is more effective in the non-perfused uterus still rich in CBG.

The results would have been completely different with ^{3}H-R5020 because, on account of its specificity, its absence of CBG binding and its much higher affinity for PR (slower dissociation of the receptor complex), progesterone can be nearly totally exchanged within 2 h at 0 °C.

^{3}H-R5020 thus measures virtually all PR and only PR[2,3].

Part of the aim of Al-Khouri and Greenstein's study was "to know the mechanism of action of progesterone in the brain". The progestin receptor measured with ^{3}H-R5020 has been shown to be implicated in progesterone activation of female sexual behaviour by comparing PR levels in conditions where progesterone is active or inactive in rats[4] and guinea pigs[5,6].

J. P. RAYNAUD
M. MOGUILEWSKY

Roussel-Uclaf,
75007 Paris Cedex,
France

1. Greenstein, B. D. & Al-Khouri, H. *Nature* **287**, 58–60 (1980).
2. Philibert, D., Ojasoo, T. & Raynaud, J. P. *Endocrinology* **101**, 1850–1861 (1977).
3. Philibert, D. & Raynaud, J. P. in *Progesterone Receptors in Normal and Neoplastic Tissues* (eds McGuire, W. L., Raynaud, J. P. & Baulieu, E. E.) 227–243 (Raven, New York, 1977).
4. Moguilewsky, M. & Raynaud, J. P. *Endocrinology* **105**, 516–522 (1979).
5. Blaustein, J. D. & Feder, H. H. *Brain Res.* **169**, 481–497 (1979).
6. Blaustein, J. D. & Feder, H. H. *Brain Res.* **177**, 489–498 (1979).

GREENSTEIN AND AL-KHOURI REPLY—There is a striking difference between the brain and the uterus in terms of the specificity of ^{3}H-progesterone binding and we feel they should be discussed separately.

As regards the brain, we showed that progesterone was far more potent than was corticosterone in displacing ^{3}H-progesterone from its binding sites in the cytosol. As corticosterone has a very high affinity for its own receptor ($K_d \sim 10^{-9}$ mol l^{-1})[1,2], one is constrained to conceive a situation in which corticosterone, but not progesterone, is denied access to the glucocorticoid receptor. Furthermore, progesterone has only one-third the affinity exhibited by corticosterone for the glucocorticoid receptor[1]. Raynaud and Moguilewsky stated that ^{3}H-progesterone did not displace endogenous progesterone from the progestin receptor, although they did not specify the tissue. It has been shown elsewhere, using ovariectomized rats, that there is rapid exchange between ^{3}H-progesterone and the unlabelled hormone at the neural progesterone receptor[3]. Whether this is a progestin receptor is by no means clear. From our results[4] the simplest conclusion is that ^{3}H-progesterone bound to a specific receptor from which it was displaced more readily by progesterone than by four other steroids.

In the uterus there is a completely different picture. We agree readily with Raynaud and Moguilewsky that there is extensive binding cross-reactivity between progesterone and corticosterone at their receptor sites. Indeed, this difference between brain and uterus is a central point of our paper. At this stage we cannot distinguish between the many components in uterine cytosol to which ^{3}H-progesterone was bound. As there is ample evidence that there is rapid and virtually complete exchange between progesterone and the tritiated steroid at 0 °C within 2 h *in vitro*, we are confident that ^{3}H-progesterone bound to progesterone receptors in cytosols from the uterus.

The substance R5020 is an interesting tool in studying the mechanism of action of progesterone, but the fact that there are moieties in brain and uterine cytosols which bind progesterone more strongly than they do R5020 militates in favour of the use of progesterone for studying progesterone receptors. In particular, if the corticosteroid-binding globulin-like material (to which R5020 does not bind) is involved in the actions of progesterone and corticosterone in the uterus, there is all the more reason to doubt whether the role of this protein will be elucidated through the use of R5020 alone.

B. D. GREENSTEIN
H. AL-KHOURI

Department of Pharmacology,
St Thomas's Hospital Medical School,
London SE1 7EH, UK

1. Grosser, B. I., Stevens, W. & Reed, D. J. *Brain Res.* **57**, 387–395 (1973).
2. De Kloet, R., Wallach, G. & McEwen, B. S. *Endocrinology* **96**, 598–609 (1975).
3. Greenstein, B. D. *J. Endocr.* **79**, 327–338 (1978).
4. Greenstein, B. D. & Al-Khouri, H. *Nature* **287**, 58–60 (1980).

Polar cap models for pulsars

ARDAVAN[1] has recently claimed on very general grounds that the emission from pulsars must be from the vicinity of the light cylinder. This seems to rule out polar cap models[2,3] in which the radiation is from particle bunches resulting from a plasma instability near the pulsar. However, the time scale with which the current distribution associated with these bunches fluctuates is much shorter than the pulsar period. A smoothed-out picture of the currents can be stationary in the rotating frame which is a reasonable starting point for theories of the magnetosphere[4]. However, such a smoothing process is unjustified in this context because radiated power varies as the square of a fluctuating current. A simple example is: N electrons moving uniformly round a circle with uncorrelated phases clearly radiate N times what each one does. No matter how large N is, we cannot calculate this radiation using the 'equivalent' steady current. In the present case, the smoothed-out current distribution still has the time dependence caused by the pulsar's rotation and hence radiates mainly from the vicinity of the light cylinder.

Because the fluctuations responsible for the radiation in polar cap models have been smoothed away by Ardavan[1], the final radiation pattern derived cannot be used to rule out these models.

RAJARAM NITYANANDA

Raman Research Institute,
Banglalore-560080, India

1. Ardavan, H. *Nature* **289**, 44–45 (1981).
2. Sturrock, P. A. *Astrophys. J.* **164**, 529–556 (1971).
3. Cheng, A. F. & Ruderman, M. A. *Astrophys. J.* **235**, 576–586 (1980).
4. Mestel, L. *Nature phys. Sci.* **233**, 149–152 (1971).

ARDAVAN REPLIES—The quasistatic condition[1] sets no restriction on the time scale of fluctuations of the electric current; it merely demands that the dependence of the current on azimuthal angle should be of the same form as its dependence on time. The arbitrary temporal fluctuations of the electric current are thus required to propagate longitudinally so as to create a rigidly rotating pattern. Indeed, unless the time scale with which the current distribution fluctuates is much shorter than the pulsar period, my analysis would predict no radiation in the range of observed frequencies: according to equations (9) and (10) of ref. 1, the power radiated into the higher order harmonics $m \gg 1$ is non-zero only when the Fourier representation of the current density as a function of $\hat{\varphi} \equiv \varphi - \omega_0 t$ contains the corresponding high-frequency amplitudes.

As long as the interparticle distances within the source do not exceed the wavelength of the radiation, the description based on smoothed-out macroscopic current densities is justified. In the example cited by Nityananda the radiation can in fact be calculated using a macroscopic description; the equivalent quasistatic current density would be

$$j_\varphi(r, z, \varphi; t) = \sum_{i=1}^{N} e\omega_0 \delta(z)\delta(r - r_0) \times \delta(\varphi - \omega_0 t - \varphi_i)$$

where the notation is the same as that of ref. 1. Furthermore, only when $N \to \infty$ does this current distribution cease to radiate[2,3], and then it does so because of the disappearance of all azimuthal fluctuations in the limit and not because of the macroscopic form or the quasistatic character of the current.

HOUSHANG ARDAVAN

Institute of Astronomy,
University of Cambridge,
Madingley Road,
Cambridge CB3 0HA, UK

1. Ardavan, H. *Nature* **289**, 44–45 (1981).
2. Jackson, J. D. *Classical Electrodynamics*, 503 (Wiley, New York, 1962).
3. Schott, G. A. *Electromagnetic Radiation*, 103 (Cambridge University Press, Cambridge, 1912).

0028-0836/81/200172—01$01.00

BOOK REVIEWS

Let them eat bread! But which bread?

John Yudkin

Medical Aspects of Dietary Fibre: A Report of the Royal College of Physicians. Pp.175. ISBN 0-272-79609-3. (Pitman Medical: 1980.) £5.50. *Medical Aspects of Dietary Fiber.* Edited by Gene A. Spiller and Ruth McPherson Kay. Pp.229. ISBN 0-306-40507-5. (Plenum Medical: 1980.) $32, £20.48.

DURING the past few years, the consumption of brown bread in Britain has increased more than threefold. It is not hard to see why this is. The Department of Health and the Health Education Council, and similar bodies in many other countries, advocate strongly the consumption of brown and wholemeal breads rather than white, and an increase in the consumption of vegetables and fruit, so as to ensure an adequate intake of dietary fibre. It is now widely accepted that brown is beautiful; and this even extends to an increased consumption of brown sugar and to an unprecedented demand for brown eggs.

Brown eggs, whatever their aesthetic attraction, have no nutritional advantage over white eggs. Brown sugar is better than white only in that a strict avoidance of the latter must lead, for most people, to a sizeable diminution in total sugar consumption. The question then remains whether there is any health benefit in eating brown or wholemeal bread rather than white.

Here are two books that summarize current knowledge, and opinion, of the effects of dietary fibre — the indigestible constituents of cereals, other vegetable foods and fruits which we used to call "roughage". The Plenum book consists of 14 chapters, by 21 different authors, describing in fair detail the research results that have been achieved in studies on the composition of fibre, patterns of consumption of fibre, patterns of consumption within and between different countries, its effects in experimental animals and in human volunteers, and its possible relation to human disease. As in all compilations of this sort, there is some overlap and repetition, as well as rather inadequate consideration of some topics. More interestingly, there are differences in the degree to which different authors are prepared to be hesitant, or enthusiastic, in accepting the largely inconclusive results of experiment as evidence of a role for dietary fibre in the prevention of human disease.

For a careful reading of the evidence shows that dietary fibre, in the quantities found in any ordinary diet, has only three certain effects that are of any consequence to human health. It can speed the passage of the gut contents, and increase the bulk of the faeces; thus, in many individuals, it will relieve constipation. It reduces the symptoms of diverticulitis, although it is often necessary to use more fibre than would normally be present in ordinary diets. Third, it reduces the absorption of some of the mineral elements, so that it may lead to deficiency, especially in the elderly.

The Plenum book, then, will be found useful by the large number of people who are nowadays engaged in studying dietary fibre. But the book would also be handy for the nutritionist or the gastroenterologist who is prepared to look in more detail at the justifications for some of the claims that have been made for dietary fibre.

The Royal College of Physicians book, on the other hand, should be read by every dietician, nutritionist and doctor. It was compiled by a committee of the RCP, whose members include some of those who wrote chapters for the Plenum volume. It covers the same ground but in the form of a simply-worded summary of the literature, with conclusions at the end of each chapter. Like the Plenum book, the diseases it considers include obesity, diabetes, ischaemic heart disease, diverticular disease and colon cancer. Here are some of its conclusions:

- ●Although societies which exist on fibre-rich diet have a low prevalence of obesity, this is not evidence that dietary fibre as such prevents obesity.
- ●No coherent pattern has emerged of the effects of fibre on the occurrence of ischaemic heart disease . . . . In view of the varying forms of fibre and their variable effects, it is not yet possible to give specific recommendations to the general public about its value in preventing heart attacks.
- ●Undoubtedly a reduction in the intake of fibre may cause constipation. Nevertheless, other factors may help to explain the commonness of constipation in Westernised communities.
- ●There are reasonable grounds for the statement that, in genetically susceptible persons, large bowel cancer could be favoured by a fibre-depleted diet, but other explanations for the commonness of this cancer in Westernised countries are possible.

But even this detached assessment of the evidence is occasionally clouded by the current all-pervading atmosphere that insists that dietary fibre plays an important role in the prevention of disease. The chapter on diabetes ends in this way: "In the long term, replacing refined carbohydrates with physically intact carbohydrate foods might also reduce the risk of diabetes by preventing or arresting obesity". Yet an earlier chapter says: "Dietary fibre has little if anything to do with preventing obesity".

With this careful appraisal of the evidence by such an authoritative body as the RCP, it becomes necessary to ask how it comes about that doctors, health educators, dieticians and nutritionists have almost with one voice been advocating an increase in the consumption of dietary fibre. The two books reviewed have themselves provided much of the answer. It is universally accepted that epidemiological study of populations with different prevalences of a particular diseases may reveal associations but not causes. Yet the major evidence still cited for the view that colonic cancer and several other diseases are caused by a lack of dietary fibre is that rural Africans are much less prone to suffer from these conditions that are urban Westerners. The difference in disease prevalence is indeed likely to be due to differences in environmental conditions, but these differences are many. Some are non-dietary, and include cigarette smoking, physical activity and competitive stress; some are dietary, and include total calories, animal protein, fat, sucrose, calcium and vitamin B_{12}, and not only dietary fibre.

In addition, there is straightforward human prejudice. Part of this derives from people's obsession with their bowels, and accounts for the belief that the many virtues attributed to dietary fibre lie in its ability to cure constipation; in 1979, they spent in the UK more than £6 million in buying non-prescriptive laxatives. Moreover, they are convinced that constipation is in turn the cause of many other and more alarming ills. With the greatest enthusiasm, then, they seized on the suggestion that all this happens because they have been eating white bread.

But the rapid acceptance of the virtues of dietary fibre is not due solely to the hope that it will relieve imaginary maladies. It is also caused — and it is still largely upheld — by the phrase "refined carbohydrate". This refers both to white flour (or bread made from it) and to white sugar. The implication is that they share similar harmful properties; on the other hand, it is believed that wholemeal flour and raw sugar share similar beneficial properties. But raw sugar is already very much "refined" compared with the sugar cane from which it is made, whereas wholemeal flower really does contain the whole of the wheat berry. Much more importantly, white flour and even pure starch have metabolic properties very different from those of sugar (sucrose).

Denis Burkitt, who, probably more than anyone else, was responsible for the revival of the view that the lack of dietary fibre is the major cause of the ills with which Western countries are plagued, ends his chapter in the Plenum book thus:

> The attitude that decries change until a case is proven is both illogical and morally indefensible. Medical history is full of instances in which effective action was taken to reduce the incidence of a disease by avoiding situations known to be associated with it, long before the cause of the disease was known or the mechanisms involved in its causation understood.

Although they write in a somewhat different context, I much prefer the statement by the authors of another chapter in the same book, as a comment on the too-hasty acceptance of epidemiological evidence unsupported by experiment:

> Such an attitude does not seem justified in view of the past continuously preached emphasis on high-fat diet as a plausible risk factor in the etiology of coronary heart disease, which has kept research in preventive cardiology at a standstill for the past 20 years.

Let us by all means eat brown or wholemeal bread if we like it. But let us not delude ourselves that it will make us healthier or prolong our lives. ☐

John Yudkin is Emeritus Professor of Nutrition at the University of London.

Incomplete expression of genes

Jay Greenberg

Cell Biology: A Comprehensive Treatise. Edited by L. Goldstein and D.M. Prescott. Vol. 3 *Gene Expression: The Production of RNAs*, pp. 684, ISBN 0-12-289503-7. Vol. 4 *Gene Expression: Translation and the Behavior of Proteins*, pp. 496, ISBN 0-12-289504-5. (Academic: 1980.) Vol. 3 $65, £36.40; Vol.4 $49, £27.60.

THESE volumes consist of a collection of articles by different authors on various aspects of gene expression. (Volumes 1 and 2 in the series dealt with cell inheritance and its molecular basis.) They represent a much-needed attempt to fill the gap between reports of original research and the conventional textbook, and they will be of greatest use to students and more advanced scientists who wish to broaden the scope of their knowledge. For the most part, the chapters are not (nor are they intended to be) in-depth, up-to-the-minute reviews for specialists. However, some of them contain detailed, well-referenced information, and they will no doubt be appreciated by experts as well as by less knowledgeable readers.

To a considerable extent these volumes succeed in fulfilling their purpose. However, they are not without shortcomings. Some of the articles are less complete, informative, scholarly and readable than others; furthermore, there is a good deal of overlap between contributions. To some extent, this is the unavoidable result of different authors writing independently of one another; however, the number of redundant passages could have been minimized by more extensive editing.

The chapters are logically organized around gene expression at the level of RNA transcription and processing (Vol. 3), and translation and post-translational behaviour of proteins (Vol. 4). Unfortunately, however, this arrangement precludes integrated discussions of systems in which gene expression has been investigated at multiple levels. Another problem is that the current rate of progress in cell biology is such that many of the chapters were out of date prior to publication. This is especially true of Vol. 3, which was published late in 1980. However, the references cited in many of the chapters do not go beyond 1978.

Probably the greatest fault (no doubt related to the preceding one) is that neither volume is truly comprehensive. Many aspects of gene expression currently of interest are barely touched. These include such general phenomena as introns, gene splicing and the cytoskeleton, as well as more specific topics such as *Drosophila* heat shock genes, immunoglobulin, globin and ovalbumin genes, and small nuclear RNAs. Control of translation in cell-free systems and virus-infected cells is discussed, as well as regulation of protein synthesis in embryogenesis (masked messengers). However, other phenomena in which translational control plays an important part are not mentioned, for example, mitosis and heat shock. There is little on gene expression in mitochondria and chloroplasts. Apparently, no additional volumes are planned even though the current series is far from complete in its treatment of cell biology. ☐

Jay Greenberg is a Staff Scientist at the Worcester Foundation for Experimental Biology, Shrewsbury, Massachusetts.

Cosmology at school

Joseph Silk

Physical Cosmology: Les Houches 1979, Session XXXII. Edited by Roger Balian, Jean Audouze and David N. Schramm. Pp.665. ISBN 0-444-85433-9. (North-Holland: 1980.) $109.75, Dfl.225.

WHAT is the cosmos made of? For all that astronomers can tell, it could be puppy dogs' tails, or back issues of *Nature*, or a host of slightly less exotic possibilities that range from snowballs to giant planets or black holes. All suggestions are equally unproven. What we do know is that the preponderance of matter in the Universe, 90 per cent or more, consists of unknown stuff, for which the term "massonium" was coined at Les Houches by the cosmologist P.J.E. Peebles. Massonium resides in galaxy halos and in the depths of intergalactic space. We infer its presence from its gravitational effect on the observable regions of galaxies. To discover the mystery of massonium is the cosmologists' dream. That it seems destined to remain so is one of the messages that is brought out in *Physical Cosmology*, the report of the 32nd session of the Les Houches Summer School in 1979.

Another elusive cosmological goal is to decide whether the Universe is closed, destined to ultimately collapse to a fiery fate or is open, to expand indefinitely to a cold and ever bleaker future. Here, the evidence presented on the one hand by Peebles and on the other hand by Gustav Tamman, Amos Yahil and Alan Sandage appears to be in conflict. Peebles cites evidence that substantial mass is present in the outer regions of galaxies and galaxy clusters, concluding that within the experimental uncertainties the Universe could be closed. On the other hand, Tamman and his colleagues measure the expansion rate out to a distance from the Milky Way galaxy of some 20 megaparsecs, and find it to be remarkably uniform. This result favours an open Universe, in which the kinetic energy of expansion greatly exceeds the gravitational potential energy that could eventually cause a contraction.

Many of the other topics in *Physical Cosmology*, which span a vast range of astrophysics, are less controversial and provide a good review of the consensus viewpoint towards cosmological issues. As might be anticipated in a lecture course arranged by nuclear astrophysicists Audouze and Schramm, cosmological nucleosynthesis receives a thorough treatment in lectures by Robert Wagoner, Yves David and Hubert Reeves. Helium is the second most abundant element in the Universe, and can be observed in a variety of astronomical objects, from the Sun and nearby HII regions to distant galaxies and remote quasars. Remarkably, astronomers find one nucleus of helium for ten of hydrogen in practically every environment that can be studied, whereas the abundances of carbon or of iron may vary by one or two orders of magnitude The big bang theory in its simplest form predicts not just the synthesis of helium when the Universe was one minute old, but even the observed abundance And this prediction is insensitive to the controversy over whether the Universe is open or closed. Cosmologists have pounced on this coincidence as evidence for the big bang theory itself

Deuterium is another element that may reflect the history of the early Universe A by-product of the synthesis of helium, deuterium is also destroyed very effectively, first almost immediately after it forms if the density of the Universe is sufficiently great, and second when stars form The high temperature achieved in the interior of a star as its nuclear fuel supply of hydrogen burns into helium results in the burning of any pre-existing deuterium By searching for matter that has not yet undergone this stellar cookery, astronomers have succeeded in detecting deuterium in very young stars and in interstellar matter The deuterium is believed to have originated along with helium in the big bang, if so, the Universe must be of low density and open in order for a significant quantity to have been synthesized

Those cosmologists who prefer a dense Universe are unhappy with this result Fortunately, the discovery announced in 1980 of the possible existence of a mass for the neutrino has managed to satisfy practically everyone Now, one has the prospect of a dense Universe dominated by neutrinos At the epoch of nucleosynthesis, the neutrinos play no role, and the Universe behaves as though it had a much lower density, corresponding to that of the protons and producing deuterium in the desired quantity The massive neutrino promises also to affect many other results in cosmology, especially the work on galaxy formation that is reviewed in the Les Houches lectures by Richard Gott, Brandon Carter and Peter Mészáros

One of the obvious omissions is the discussion of the pancake theory of galaxy formation The Moscow school of cosmologists is its most ardent protagonist Great structures form on the scales of clusters of galaxies, and provide an intriguing picture of the Universe as it appears to be on very large scales, full of sheet-like arrays, filaments and gaping holes. But on the largest scales of all, on the order of and beyond the present horizon, homogeneity is very much apparent, as the lectures of Martin Rees bring out

One of the joys of *Physical Cosmology* is its illustrations of the lecturers in action From Wagoner on the grass to Rees at the blackboard, these candid shots add a dash of vitality that provides an extra reward for those brave souls who have mortgaged their possessions to buy this book □

Joseph Silk is Professor of Astronomy at the University of California, Berkeley

---

Sites of special spiritual importance

Kenneth Mellanby

The IBP Survey of Conservation Sites. An Experimental Study IBP, 24 Edited by A R Clapham Pp.344 ISBN 0-521-22697-X (Cambridge University Press. 1980.) £25, $57 50

ONLY within the past hundred years have there been attempts to conserve wildlife for its own sake, and to establish nature reserves in which plants and animals can continue to exist under something approaching natural conditions It has not always been easy to decide upon what criteria nature reserves, national parks and other protected areas should be selected In the early days we relied on the instincts of a few dedicated naturalists. Today, many would like to use more scientific methods.

The International Biological Programme (IBP) was established in 1964, to try to obtain a better understanding of the environment as a basis for rational management of natural resources Much of its work concerned productivity and potential food production in a world with a growing population Its Terrestrial Conservation Section (CT) decided "to initiate a world-wide study of the great variety of terrestrial ecosystems, leading to the selection of representative samples of all the more important ones" This book contains the results of the survey

The method used to evaluate sites was the IBP check-sheet survey 13,000 copies of the check-sheet, details of which are given in a 22-page appendix, were distributed world-wide 3,010 were returned, 2,690 satisfactorily completed Most returns came from the developed countries, Britain, the USA, Europe (where Poland excelled), Japan and Australia These results may seem disappointing, but they gave much more information than had previously been available on a world scale. The book describes how the information was processed, and includes critical discussions of the methodology It should be of great value to those concerned with the survey and classification of vegetation on an international and a national scale With its appendices dealing with different techniques for vegetational classification it will be a useful tool for others involved in this field, though it can hardly be recommended as a good read even for the most dedicated ecologist

No one can doubt the scientific value of this study, as an academic exercise in ecology Its value in the selection of potential nature reserves is more doubtful In the past most reserves have been established because the area has appealed to the emotions of the selector, and because it has been available. The process has been essentially hit or miss — it may be surprising that there have been so many hits, and that so many of the older reserves are still obviously of considerable value There are many who still consider that this is the best method of selection, particularly in Britain where most of the treasured sites have been profoundly altered by human activities Many of the scientific reasons for nature conservation, such as maintenance of a large heterogeneous gene-pool or the perpetuation of the full diversity of the world's plant and animal communities, have been described as rationalizations of decisions made for aesthetic and emotional reasons. It has even been suggested that Sites of Special Scientific Interest might be better designated Sites of Special Spiritual Importance.

I believe that these subjective emotions are often the reason why many of us are interested in wildlife conservation However, this does not mean that a scientific study of the results of emotions is not still valuable. So this exercise is to be welcomed as giving us a clearer picture of the world scene, and if its application means that more unspoiled areas can be preserved for future generations we should all be grateful for those who completed what, to me, would have been a very tedious operation □

Kenneth Mellanby is President of the Bedfordshire and Huntingdonshire Naturalist Trust His latest book, Farming and Wildlife, *will be published by Collins on June 14*

Polyhedra for chemists (and aesthetes)

Michael Mingos

Transition Metal Clusters Edited by B F.G Johnson Pp 681 ISBN 0-471-27817-3 (Wiley 1980) £33, $88 15

THE scientific interest in polyhedra can be traced back to ancient Greece, and the understanding of the mathematical properties of polyhedra developed then and subsequently have often proved to be relevant to physics and chemistry With this long tradition, therefore, it is not surprising that during the past 20 years transition metal chemists have been particularly interested in synthesizing new "cluster" compounds which have between four and forty metal atoms bonded together in a wide variety of polyhedral arrangements Aside from the aesthetic satisfaction of elucidating the structures of these beautiful molecules, such compounds have provided a range of interesting theoretical valency problems and have stretched to the limits the techniques currently available to chemists for characterizing molecules. In addition, they have stimulated the interest of industrial chemists and molecular biologists, the former because of the potential of the compounds as homogeneous and heterogeneous catalysts, and the latter because iron sulphur clusters have been found to be important components of ferredoxin and nitrogenase.

Transition Metal Clusters therefore represents a timely attempt to bring together under one cover some of the developments in this area.

The book covers most of the important aspects of cluster compounds, with P. R. Raithby reviewing the geometries of more than 1,000 cluster compounds whose structures have been elucidated using X-ray and neutron diffraction techniques. K. Wade and R. G. Wooley have contributed two very different views of the bonding in these polyhedral molecules, J. A. Connor has provided a balanced and critical account of the thermochemical estimation of metal–metal bond strengths in cluster compounds, and B. F. G. Johnson and R. E. Benfield have discussed their fluxional characteristics. The biological significance of cluster compounds is stressed by C. D. Garner, in a chapter dealing with cubane clusters, and R. Whyman has attempted to evaluate whether cluster compounds do actually function as effective catalysts by reviewing the relevant academic and patent literature. There is also a detailed review by A. J. Deeming on the reactions of metal carbonyl cluster compounds.

The book gains by drawing on the expertise of the authors to provide a reasonably up-to-date (literature references up to the end of 1978) and broad view of the field, but suffers from an uneven view of the potential readership of the book. Thus although some chapters are written at a level which would be appreciated by a third-year undergraduate, others are in the style of a research paper or a terse contribution to a Chemical Society *Specialist Periodical Report*. Therefore, although I wholeheartedly agree with the editor that this book will serve as a "source of reference to workers in the field", I very much doubt if it will be used much as "an advanced text for undergraduates or graduates in chemistry". □

Michael Mingos is a Lecturer in Inorganic Chemistry and Fellow of Keble College at the University of Oxford. He was recently awarded the Corday Morgan Medal of the Chemical Society for his experimental and theoretical contributions to organometallic and cluster chemistry.

Animals: their use and their abuse

Michael F.W. Festing

Animal Suffering: The Science of Animal Welfare. By Marian Stamp Dawkins. Pp. 149. Hbk ISBN 0-412-22580-8, pbk ISBN 0-412-22590-5. (Chapman & Hall: 1980.) Hbk £7.50, pbk £3.95.

ON AVERAGE, a person living in a developed country such as the UK will be responsible for the death of about 600 chickens, 22 pigs, 20 sheep and 7 cows for food and 5 mice, 2 rats, 1 chicken, a third of a guinea-pig and two-hundredths of a cat and dog for medical research during their 75-year life span. They will also be responsible for the deaths of countless fish, both for food and for sport. Their food supplies and health will also be maintained at the cost of the death by poison or trapping of an unknown, but large number of mammals and birds classified as pests. Who can doubt that the production, maintenance and killing of all these animals involves a certain amount of suffering? The problem for us as individuals, and as a society, is to decide what to do about it, if anything.

In order to make some progress, we must sort out our moral attitudes to animals and to our fellow human beings, and we must attempt to strike a balance between the amount of suffering which we cause and the benefits which we derive from the animals. This is an extremely difficult problem, because it is almost impossible to quantify either the benefits of, say, using animals in medical research, or the amount of suffering that we cause. This book is an attempt to summarize our knowledge about suffering in animals, and also point out many inconsistencies in our attitudes to other species.

Following an introductory chapter, the assessment of animal suffering is considered from several different points of view. These include suffering and health, "productivity", the "unnatural" life, the physiology of suffering, behaviour, what animals choose and analogies with ourselves.

Physical health is seen to be a necessary, but certainly not a sufficient condition for freedom from suffering, and is probably the single most useful criterion of animal welfare. Productivity, in the sense used in agriculture, is more concerned with economics than animal welfare. A higher productivity (i.e. profit) may well be achieved by putting an additional chicken into a battery cage, but this does not reduce the suffering of those already in occupation. In fact, reading these chapters it becomes clear that there is no single way in which the suffering of an animal can be clearly quantified. Ideally, it would be nice to be able to measure some physiological or biochemical variable which could be used as an index of "stress". Unfortunately, although certain parameters can be measured which give an indication of the state of the general adaptation syndrome, there is no evidence that a complete absence of "stress" (which can itself be defined in many different ways) is necessarily beneficial to an animal. Nor can we say that an animal suffers because it is forced to lead an "unnatural" life. In nature many animals have a relatively short life-span, quickly succumbing to predators, disease, exposure or starvation. Their death may be lingering and painful. Even allowing animals to choose the conditions under which they will live does not always help. Animals sometimes choose a course of action which is not in their best interests; few domestic pets enjoy a visit to the vet, even though this is usually for their benefit.

Does this mean that we are unable to assess animal suffering at all? Dr Dawkins is clearly concerned that the general tone of these early chapters "may have seemed rather negative", and she goes on to provide a list of questions which we must ask ourselves in order to determine whether an animal is suffering unduly. We must consider the conditions in which the animals are kept, whether they are healthy, whether they appear physiologically or behaviourally abnormal, and if so what is likely to be the cause of these abnormalities. If possible we should also determine what conditions the animals themselves prefer. And in the end, the assessment will have to be subjective. "There is no formula for giving a clear yes/no answer to the question of whether an animal is suffering." Even if we could answer the question unequivocally, we would still have to decide how much suffering we should permit in relation to the benefits we obtain from animals. That, however, is a whole new subject which falls more in the field of ethics than biology.

Dr Dawkins has produced an excellent review of the state of our ignorance about suffering in animals, and has highlighted many inconsistencies in our attitudes. This book should be read by all those involved with farm and laboratory animals, in the hope that they will endeavour to reduce needless suffering wherever it occurs. If we must cause suffering to animals in the name of agriculture or medical research, let us at least make sure that it is not caused unnecessarily through our ignorance of the needs of those creatures that provide us with so many benefits. □

Michael F. W. Festing is at the MRC Laboratory Animals Centre, Carshalton, Surrey.

nature

Pālit Chemistry Library
Science College
92 Upper Circular Rd
Calcutta

15/2/82

21 May 1981

British universities must win lost friends

For British universities, the crisis of the past several years is about to reach another climax. Last week, the University Grants Committee sent off its promised letter, which emerged as a restatement of a now familiar problem. Next month, universities will be told individually what they will be able to count on from central government in the next three years. Then, with this academic year almost over, universities will know whether the prospect for the immediate future is a bad as they have been fearing these past several months, or not quite as bad. There is no chance that they will be pleasantly surprised, although it will be of some comfort that the government has now promised that the misery ordained to begin this summer will at least be assured for a full three years. At the end of that time, the British university system will be transformed. Almost certainly, it will be smaller. Nobody knows quite how.

The nature of the impending change is by now well understood. From the beginning of this academic year, the central government's support for universities has been decreased by the notional cost of educating overseas students, which the universities are expected to recover by charging economic fees. If the market for British higher education were insensitive to price, the financial consequences would be negligible. In practice, since unknown proportions of overseas students will now go elsewhere, a substantial amount of money will be taken out of the system. The consequence may be a reduction of up to seven per cent in the aggregate income of the universities, but the consequences will not be uniformly spread. Some universities will fare worse than others, but to a degree quite unrelated to their intrinsic worth. The British government has also decided that there should be a reduction in the resources made available to the universities. This will amount to just over 3 per cent next year, and to 8.5 per cent in three years. Throughout the university system, committees have been working hard, making lists of the steps they may have to take in the months ahead. Getting rid of academic departments, and of people, appears on most prospectuses. Shutting down whole universities has been canvassed as a possibility, but is unlikely.

Many universities and university teachers are bewildered by the abruptness of the change that has overtaken them. Gone for good are the happy 1960s, when both Conservative and Labour governments urged expansion (so as to increase the proportion of young people going on to higher education) and appeared willing to provide the means. What can have changed? And why has the change come so suddenly? Many academics are still plaintively asking these questions. The answer is that the change is not nearly as abrupt as it may seem. The climate has been changing since the early 1970s, since when public support for the universities and their research has stagnated. For most of that time, academics have been foolishly complacent, often behaving as if the deterioration of their institutions was the responsibility only of the central government, aided and abetted by the University Grants Committee. Now, the government can fall back on the thin excuse that further impoverishment is justified by the impending reduction of the age groups from which university students are recruited. The underlying truth, however, is that the British universities have lost their political friends. Many in the House of Commons will be sympathetic to the plight of the universities in the next few months, but few elected politicians would use up political capital by defending the universities from the cuts ahead.

The universities are not blameless. From time to time, they have almost courted unpopularity. Student troubles have not helped, even though the British university has been less threatened by them than many others. Even now, however, minor issues concerning students affect the capacity of the universities to keep their political friends. Is it right, for example, that the cost of running students' unions, hitherto a charge on local authorities, should now be paid from the dwindling budget of the University Grants Committee? Why should not students pay? The more serious alienation of the universities from their traditional supporters in the House of Commons springs more directly from the assumption of the universities that even in times of rapid and painful social change, the structure of the university system itself should remain untouched, changing only if the universities so decide. One result is that even those outside the system who value academic freedom are asking whether academic freedom necessarily requires lifelong tenure for all. Then those among the universities' natural friends who accept that scholarship is an indispensable part of the institution of the university are unconvinced that the universities are vigilant enough in distinguishing good scholarship from the less good or even the downright bad. The universities have lost many potential friends by their unimaginative, even unhelpful, responses to several attempts to reform the transition in the British system from school to university. And their proper insistence that universities are self-governing institutions is weakened by their frequent demonstrable failure to manage their own affairs effectively.

Hard words such as these are not inappropriate even in hard times. Indeed, if the universities do not now acknowledge that they are partly responsible for the plight in which they find themselves, they will be less able to meet the crisis the British government has unthinkingly forced on them. They will, in particular, be less able to make the plain argument that the policy of retrenchment being forced on them is a foolish policy, innocent of clear objectives, marred by vindictiveness and almost calculated to ensure the British taxpayers get even less value for the diminished sums of their money to be spent on higher education in the next few years. For several years, it has been clear that the most urgent need in British higher education is for a reconsideration of what it is for, how it should be financed and how different kinds of institutions, universities included, should contribute. Probably it is now too late for that reappraisal to be carried out deliberately. But radical thinking by threatened institutions might still do wonders.

It would be a useful start if the universities would decide where they stand in relationship to each other and to comparable institutions, the polytechnics for example. In the past few months, it has been fashionable among the enemies of higher education to say that there are too many universities. That assertion is unverifiable, and probably untrue. But there are too many universities of the same kind, or too many universities striving after an idealized vision of what a university should be — perhaps a cross between Oxbridge and the Harvard Business School. All except Scottish universities insist that, for most students, three years of undergraduate teaching is sufficient for a first degree. All insist that scholarship is inseparable from teaching and, more literally, that there must be a full range of research programmes undertaken on the premises. Most of them are so balkanized internally, with power and responsibility delegated to separate departments and colleges, that academic policy is a kind of statistical amalgam of different people's inclinations. And most universities scorn most other kinds of higher education than their own — the British polytechnics or

0028 0836/81/210177-02$01.00

liberal arts colleges on the pattern of the United States

Breaking this log-jam of prejudice against change is desirable in itself but could also have practical consequences in the present crisis Given the quality of their students and the scale of their research programmes, many universities would be better off if they could become liberal arts colleges in the best sense — providing students with a general education (which includes an education in science) without the encumbrance of in-house research programmes (which does not imply that academics need not be scholars) Other universities, by contrast, might make their way more successfully in the world by concentrating more vigorously on the kind of vocational education that has made institutions such as the Cranfield Institute of Technology (not strictly a part of the university system because it is financed directly from central government) both successful and valuable British universities will be the more able to turn the tight corner ahead if they come to believe that diversity within the system should be an objective — and if the University Grants Committee can devise an incentive for helping them stay with such objectives

Hotchpotch

The hotchpotch system of financing universities also needs reform Money for day to day purposes reaches university treasurers from three principal sources — from the Department of Education and Science on the recommendation of the University Grants Committee, from local education authorities which pay the fees of all home students (and which are reimbursed by the Department of the Environment) and from individual students from overseas, who must pay their own tuition fees as they go along But the local authorities are also required by law to provide many students with a maintenance grant calculated after allowing for parental income Most universities have only a negligible income to call their own, although many of the Oxbridge colleges are still well-heeled In most universities, tuition fees now exceed twenty per cent of direct subvention while maintenance grants paid from public funds to university students are probably close on forty per cent of direct income The government seems for the time being to have given up its ambition to replace maintenance grants by loans, but the impending crisis is caused largely by its threatened reduction of the direct grant But in this crisis, there must be room for a redistribution of these substantial funds When the interest of students in their education is matched by the universities' anxiety to survive, however, there must be places where it would be in everybody's interest — even the government's — that funds were transferred from the maintenance grant to the general budgets of the universities The simple way of doing this would be to require that students should work for the institutions that educate them, thus reducing general costs British students might in the process lose their unique status as perpetual dependants, but the world would not thereby end

Hardship

Academics must also be prepared for hardship The impending crisis has already raised questions about security of tenure — and many universities are dangling the prospect of early retirement before their staffs In the period ahead, however, the universities may not be able to afford this dubious luxury — and the government is unlikely to help So academics may have to make do with less pay, at least in real terms Will they? The Association of University Teachers is already up in arms at the cuts in prospect, and is holding out the hollow threat of a strike The academic profession will, however, be further discredited if, after decades of high-sounding (and cogent) argument in defence of the university as a social institution, academics let universities collapse by insisting that national agreements on pay must apply to everybody Indeed, they should go further than be flexible on pay and agree to be flexible about the work they do Research, they rightly insist, is inseparable from the university, but so is teaching And those who do less research must be prepared to do more teaching

Such developments, even when designed to meet the immediate crisis, cannot easily be brought about Few British universities have mechanisms by which their academic members can agree on the allocation of car parking space, let alone on matters even of academic policy Yet academics are understandably suspicious of ceding such responsibilities to the university administrators So even when the going is good, too little is decided At times like the present, when ingrained suspicions are heightened, paralysis tends to supervene Yet the crisis requires that universities should be individually equipped to make decisions of two sharply different kinds

First, there will have to be decisions about the academic qualities of individual academics Whose research is worth continued support? Who might fairly take on more teaching? Who may have to be asked to leave? And which departments are more dispensable than others? Traditionally, academics have avoided such decisions because they are invidious If they do so now, decisions will nevertheless be made but in an unseemly way — or by drawing lots And the universities as well as unlucky academics will be the losers

Second, universities need ways of deciding strategies for survival Committees of academics are ill-equipped for such tasks, or even comprehensively to list the alternatives (Courses of action unpalatable to some are rarely heard of) The standard compromise is to appoint a committee with an external and supposedly independent chairman to suggest what might be done — the mechanism is now being followed by the University of London The trouble is that the recommendations that follow have no force, and can be endlessly disputed As things have turned out, there is no choice for the universities but to agree that their administrators have a duty to draw up the agenda for the inevitable debate on what course should be followed — and also a right to bring an end to argument In reaching what must be painful decisions, the universities singly would benefit from some means of knowing that they are not all rushing off in the same direction but, rather, moving towards diversity

Self-interest

The peculiar difficulty of the problems with which the universities are now confronted is that their own narrow self-interest may not coincide with the long-term interests of the university system as a whole or with the interests of their customers — in the last resort, their potential students and those prepared to sponsor students, ultimately the taxpayers Plainly, however, it will serve no useful purpose if individual universities decide to concentrate on courses for which there is no demand, or on fields of research for which there are no sponsors Up to a point, such market forces will keep individual universities moving in prudent directions But the market for higher education is far from perfect — and should in everybody's interest remain so It would be tragic, and a travesty of the idea of a university, if students' choices of courses were determined exclusively by short-term economic considerations, or if grant-making bodies made grants for research on the basis of expected communal benefit Both students and the sponsors of research are known to make mistakes And the social value of universities derives directly from their capacity to help devise an uncertain future

These are the steps that British universities must take in the weeks immediately ahead There is another course of action that they cannot neglect The government's decision that there must be a cut in university funds has been plucked like a rabbit from a hat The consequences have not apparently been considered except by those among the government's economists concerned with the size of the borrowing requirement So far, both the University Grants Committee and the universities have been willing to administer this sudden change of policy without any but the mildest protests Lacking friends, the universities may think they have no choice But, in the long run, the present crisis will seriously damage the interests of a large part of the British population It is time that the universities abandoned good manners, and told the government what will be the consequences of its action And it is not too soon to set out to win back the lost constituency of friends, presumably first among parents whose children will, on present form, be deprived of higher education

0028 0836/81/210178 01$01 00

Creationism again an issue in Arkansas

Civil liberties take suit against law

Little Rock, Arkansas

The American Civil Liberties Union (ACLU) has decided to engage in a full-scale confrontation with the growing creationist movement over the teaching of evolution in US public schools. Early next month, ACLU is planning to file a complaint against the State of Arkansas arguing that a law passed by the state legislature in March — which requires that "creation-science" be given equal time with the theory of evolution whenever the latter is discussed in public school classrooms — is unconstitutional, since it ignores the required separation of church and state.

The ACLU case may well find its way up to the US Supreme Court. It seems certain to generate as much controversy as the last time that the organization confronted the creationist arguments. This was in 1928 when Clarence Darrow, acting as an ACLU volunteer attorney, defended Tennessee biology teacher John Scopes against the charge that he had violated state laws by teaching evolutionary theory in the classroom.

Although Scopes was convicted and fined $100, Darrow won a moral victory — and, over the following decades, individual states removed from their statute books laws which prevented Darwinian evolution from being taught.

This time, however, the tables have been turned. The creationists have spent many years developing the argument that their ideas can be described as "scientific", and in eliminating all explicit religious references from public school textbooks promoting the idea that the world originated in the hands of a divine creator. More recently, they have spent an equal amount of legal effort in eliminating all explicit religious references from proposed legislation at both the state and the federal level, demanding equal time for what are described as two rival scientific theories.

The legislation introduced in Arkansas, for example, has been carefully crafted by a creationist group in South Carolina, with advice from sympathetic lawyers across the country, to answer precisely the type of challenge which ACLU is planning to mount.

Various other states have been discussing similar "equal time" bills over the past two years. Last year, for example, the Georgia legislature just failed to approve a bill in the closing minutes of its 1980 session. And according to Mr Paul Ellwanger, head of the group in Anderson, South Carolina, which wrote the Arkansas bill (known as the Citizens for Fairness in Education), seventeen other states have been considering similar or identical bills in their current or recently-ended sessions.

Mr Ellwanger emphasizes that his own group is not formally linked to any political or religious organization. In Arkansas, however, two significant factors seem to have played a major role in getting the "equal time" legislation on to the statute book. The first was the important role played by groups such as the state branch of the Moral Majority and the Little-Rock-based organization Family, Life, America, God (FLAG) in stimulating support for conservative candidates — such as Arkansas' current governor, Mr Frank White — in last year's election. The second was the speed with which state lawmakers were prepared to push the bill through both houses of the legislature.

The bill was introduced into the Senate by Mr Jim Holsted, a first-term Senator from North Little Rock, a wealthy suburb of the state capital. He has admitted that he knows little about creation-science, but as a "born-again" Christian is a firm believer in the book of Genesis. According to Holsted, he was given the bill in virtually completed form by a mathematics teacher at a Jacksonville High School, who had already been attempting to persuade the local school board to require "equal time" for creationism and evolution in the schools under its jurisdiction (a tactic which has already been successful in several states).

The bill was passed overwhelmingly on the floor of the state Senate. Several members of the Senate said later that, although they did not necessarily agree with the content of the bill, they had been reluctant to vote against it, since to have done so would have been interpreted as a vote against God — a heavy risk to take in a strongly Bible-oriented state.

The bill was introduced almost immediately into the House of Representatives by a business neighbour of Mr Holsted's, Representative Cliff Hoofman. It passed to the education committee, where the bill's supporters and opponents were each given five minutes to put their case — the sum total of public discussion of the bill's contents. A motion to delay the committee vote was refused by the

Arkansas legislation reflects old quarrels

The Arkansas law, passed in March (and introduced in virtually identical versions in seventeen other states), is known as the "Balanced Treatment for Creation-Science and Evolution-Science Act". Its principal requirement is that "public schools within this state shall give balanced treatment to creation-science and to evolution-science. Balanced treatment of these two models shall be given in classroom lectures and in textbook materials taken as a whole for each course, and in library materials and other educational programs".

"Creation-science" is defined as the scientific evidence for creation and inferences from that evidence. This includes evidence of: sudden creation of the Universe, energy and life from nothing; the insufficiency of mutation and natural selection in bringing about development of all living kinds from a single organism; changes only within fixed limits of originally-created kinds of plants and animals; separate ancestry for man and apes; explanation of the Earth's geology by catastrophism, including the occurrence of world-wide flood; and a relatively recent inception of the Earth and living kinds.

The bill contrasts this with the teaching of "evolution-science" — referred to earlier as the theory of evolution — and the evidence which is held to support it. Whereas previously both creation and evolution had been referred to as "theories" to which equal time should be given in the science classroom, the bill passed in Arkansas refers to the two "scientific models" of evolution-science and creation-science; it also states as a "finding of fact" that "only evolution-science is presented to students in virtually all of those courses that discuss the theory of origins. Public schools generally censor creation-science and evidence contrary to evolution".

Particularly controversial has been a passage in the bill which states that one of its purposes is to prevent "establishment of theologically liberal, humanist, nontheist or atheist religions". The bill goes on to state that public school presentation of "evolution-science" is unconstitutional, because "it produces hostility toward many theistic religions and brings preference to technological liberalism, humanism, nontheistic religions and atheism, in that these religious faiths generally (*sic*) include a religious belief in evolution".

The Arkansas law also states, in a key passage designed to provide protection against the anticipated challenge, that: "public school presentation of both evolution-science and creation-science would not violate the constitution's prohibition against establishment of religion, because it would involve presentation of the scientific evidences and related inferences for each model rather than any religious instruction".

0028-0836/81/210179-02$01.00

chairman, and after a short discussion it was passed and sent onto the floor of the House, where it was brought up and hurriedly passed in the closing hours of the 1981 session at the end of March.

Governor White lost no time in signing the bill, admitting later that he had not read it before doing so. His decision provoked a storm of protest from scientists in the state, most of whom were unaware of the bill until it had been passed into law. The Arkansas Academy of Sciences at its annual meeting passed a resolution urging that the law be rescinded, as did 300 faculty members of the University of Arkansas at Fayetteville, who complained that the act would hurt the university's reputation.

Governor White, however, has made it clear that he does not intend to take any steps to repeal the law unless it is judged as unconstitutional. And many Little Rock observers feel that he is likely to stick to this position, given the extent to which his surprise victory over incumbent Democrat Governor Bill Clinton last November seems to have been helped by the active involvement of groups such as FLAG and the Moral Majority in encouraging the conservative vote.

As a result, the spotlight is now focused on the ACLU complaint, which both sides see as setting important precedents. Groups which are likely to support ACLU include the National Association of Biology Teachers, which has long campaigned about allowing creationist theories to be placed on the same level as evolution in school biology classes. And the President of the Arkansas Academy of Sciences, Professor John Kenneth Beadles, head of the State University of Arkansas' biology department, is suggesting to fellow academy members that they add their support to the ACLU effort.

The creationists are getting ready to meet the challenge. One of the principal advisers on the drafting of the Arkansas Bill is said to have been Wendell Bird, an attorney who was at one time an editor of the *Yale University Law Review,* and who now works with the Institute for Creation Research in San Diego, California. The institute, as a tax-exempt organization, takes pains not to be seen to be involved in legislative campaigns; however, Mr Bird's writings are referred to by Mr Ellwanger and others as powerful support for the argument that the teaching of creation-science can be defended by referring to rights guaranteed by the constitution.

If ACLU wins its case, either at the district court level, during the appeals process, or in the Supreme Court, it will be a major setback not only for the creationist movement — which now claims over 750 scientists among its ranks — but also for the broader poltitical strategy of the conservative right. If it loses, it will be a significant boost to creationists, who are already predicting that bills similar to that passed in Arkansas are likely to be introduced in 40 state assemblies next year.

Pressure is also building up at a federal level. Several Congressman have received a model bill promoting research support for "creation-science", from a group, also headed by Mr Ellwanger, called the Citizens Against Federal Establishment of Evolutionary Dogma. It seems only a matter of time before some congressman introduces the legislation.

The two sides are as far apart as ever. Mr Ellwanger argues that "fairness should not be an option; we now have the tools to insist that fairness be mandated by law, in the interests of academic honesty". The Arkansas *Gazette*, which has carried out an uncompromising campaign against the bill, warns that it will brand the state as a "bastion of know-nothingness in the teaching of science." The court case is unlikely to stay polite.

David Dickson

Scopes again?

The present is not the first occasion when Arkansas has been the focus of national attention on the constitutionality of efforts to influence the teaching of evolution in public schools. In 1968 the US Supreme Court struck down a law, introduced into the state in 1928, which made it illegal for any teacher and educational institution supported by state funds to teach "the theory or doctrine that mankind ascended from a lower order of animals".

Following the Scopes trial, no serious attempts seem to have been made to enforce this law. However, the case was passed to the Supreme Court after the statute had been unsuccessfully challenged in the state's Supreme Court by a biology teacher in Little Rock's Central High School, Mrs John O. Epperson.

In declaring the state law unconstitutional, Justice Abe Fortas argued that the 1928 law, enacted in a popular initiative, was a clear effort inspired by the Scopes trial to outlaw all but the "Bible story" of creation. "There can be no doubt that Arkansas has sought to prevent its teachers from discussing the theory of evolution because it is contrary to the belief of some that the Book of Genesis must be the exclusive source of doctrine as the origin of man" he said.

At the time of the trial, Arkansas and Mississippi were the only states to maintain an anti-evolution law. The fact that the law preventing the teaching of evolution was judged by the Supreme Court to infringe the "establishment of religion" clause in the constitution is expected to be used by the American Civil Liberties Union in its attack on the new law. Scopes himself, then living in Louisiana, hailed the Epperson decision as "what I have been working for all along". His own conviction was quashed by an appeals court, so he was unable to challenge the constitutionality of the Tennessee law.

Sakharov conference

Birthday honour

New York

The two-day conference at Rockefeller University on 1–2 May in honour of Dr Andrei Sakharov was an ironic occasion. The opening day was celebrated in the Soviet Union with the customary package of approved slogans, the forty-first of which reads "Soviet scientists! Improve the effectiveness of research! May the alliance between creative thought and creative labour be much stronger!" Sakharov meanwhile remained in internal exile in Gor'kii.

The conference had been called in part to mark Sakharov's birthday on 21 May and consisted of three parts — Sakharov's work in science, for nuclear disarmament and for human rights. Among the scientific papers was one by John Archibald Wheeler (University of Texas) on Sakharov's "sausage-skin" cosmology, another by Harold Furth (Princeton) on Sakharov's contributions to fusion research and a third by Val Fitch (Princeton) on anti-matter asymmetry. Those present nostalgically recalled Sakharov's involvement with the international scientific community in the past twenty years and argued that his present isolation was tantamount to bringing his scientific career to an end.

The conference was at something of a loss to reconcile Sakharov's work on the Soviet thermonuclear bomb in the 1950s with his more recent concern for disarmament; the argument that by assisting in the strategic parity between East and West he was helping to satisfy one of the realistic preconditions of disarmament was regarded by some as artificial and by others as unnecessary.

Sakharov's own message to the conference (see page 184) was largely concerned with human rights and, in particular, with the plight of Soviet citizens other than himself. On the question of the relationship between professional scientists in the West and their Soviet colleagues, Antonino Zichichi, president of the European Physical Society, told the conference that more than a thousand scientists had undertaken that until the end of this year, all scientific papers submitted for publication would bear the legend "In honour of the 60th birthday of Andrei Dmitrevich Sakharov". For the time being, however, the disputed question of a scientific boycott of the Soviet Union such as that organized last year, although aired at the conference, has been allowed to lapse.

Another of the ironies of the timing of this month's conference is that it followed soon after the celebration of the twenty-fifth anniversary of the Joint Nuclear Research Institute at Dubna, south of Moscow. In a full-page feature in the daily newspaper *Sotsialisticheskaya Industriya,* attention was drawn to the search for

0028-0836/81/210180-02$01.00

transuranic elements in the "island of stability" (atomic numbers 114–120), studies of black holes and quarks. Mr Leonid Brezhnev, at the twenty-sixth Party congress, had emphasized the need for "fruitful cooperation of all states" on these problems, yet Sakharov has been unable to meet or correspond with his colleagues abroad since 22 January 1980.

Vera Rich

UK industrial research

All change

Britain's Department of Industry has reorganized the way in which it supports industrial research, seemingly to counter criticism that it has been poor at defining priorities and has relied too heavily on reacting to research proposals from industry. The nine research and development requirements boards, each aimed at a particular sector of industry, have been replaced by five boards, each dealing with a particular technology.

The five new boards are to cover the same spread of research as the old ones, the idea being that by increasing the spread of each board's responsibilities the government's definition of long-term objectives and priorities will be made easier. Hence for example, the new Mechanical and Electrical Engineering board will take over the work of the old Ship and Marine, Mechanical Engineering and Machine Tools and Electrical Technology boards. The old Chief Scientist's board, which considered topics not covered by the other boards, is to be incorporated into the new Textile and Other Manufactures board.

The other new boards are Electronics and Avionics which replaces Computer Systems and Electronics, Materials and Chemicals which replaces Chemicals and Minerals and Engineering Materials, and Metrology and Standards which remains the same under the chairmanship of Dr Duncan Davies, now called Chief Engineer and Scientist in the department.

The functions of the new boards will remain largely unchanged. Industry and universities will be able to apply to them for government support for industrial research and they will continue to place contracts for research of their own choosing within industry and the department's own research establishments. The arrangement whereby levies are charged on the profits accruing from research funded by more than 50 per cent by the department will remain. And each board will have an executive committee to approve projects. It is hoped that the reorganization will encourage the boards to place more contracts in priority areas of research. However, the fact that there are no plans to appoint subcommittees has led to some fears that the boards' remits could be too wide to be effective.

The department's budget for industrial research is to remain roughly constant. This year it is £600 million, £35 million of which will be spent in industry and £25 million in the department's research establishments. A small saving in funds will be made from the reduction in staff expected to follow the reorganization.

Judy Redfearn

Bioengineering hazards

Europe doubts

Brussels

An inconclusive meeting on the safety aspects of research on recombinant DNA was held here last week by the European Community's Economic and Social Committee (ESC). At this stage, it is hard to tell whether the meeting will influence the development of Community legislation on the subject — and, if so, in what direction that influence will be directed.

The discussions ranged from the long-term ethical problems posed by the possibility of tampering with human genes to the risks of accidentally releasing a new and dangerous strain of microorganism. The impetus for the meeting was the European Commission's draft recommendation on the registration of DNA research — ESC is required to give an opinion before the Council of Ministers make a decision.

Most participants seemed to agree that industry was already being over-cautious. By trying to avoid the mistakes of the nuclear industry, the safety measures recommended in the early 1970s for DNA research now seem to reflect exaggerated fears.

One pointer in this direction was the conclusion by the European Science Foundation's Liaison Committee on Recombinant DNA in January that recombinant DNA research as such entails no significant novel biohazards. It was also stressed that recombinant DNA techniques can be safer than conventional methods for studying pathogens and toxins. Safety guidelines for handling pathogenic viruses, it was argued, were far more necessary, and the World Heath Organization's guidelines are late in arriving. These are due to be published within the next few months.

It was over the social risks that most of the participants were divided. Representatives of trades unions pointed out that there was a danger that decision-making would be restricted to industry. British speakers pointed out that the British Genetic Manipulation Advisory Group has retained control, but French participants conceded that this was not the case in France and elsewhere. Several participants were convinced that the concentration of knowledge in the hands of industry will lead to commercial (or even military) applications not necessarily beneficial. Calls for less secrecy were, however, answered by reference to the fact that all patents are by definition publicly available.

But does the public have sufficient say over the uses to which the recombinant DNA techniques will be put? This issue overlapped with that of the role of the media in informing the public. By attempting to bridge the gap between the scientist and the layman, was there a danger that the press was exaggerating the dangers or speed of developments in this field?

Although many of the speakers argued that the DNA guidelines were safety regulations looking for a hazard, some members of ESC confirmed afterwards that they were in favour of legally enforced safety regulations, even if they needed to be frequently revised to take into account changes in risk assessment. It was suggested that the Community was merely going through the same process that had taken place in the United States a few years ago, but a little later because of the several countries involved.

Jasper Becker

US defence budget

Laser destruction

Washington

The US Senate last week proposed adding an extra $50 million to the budget of the Department of Defense to speed up the research and development of a space-based laser system designed to shoot down incoming enemy missiles. At the same time, a Senate appropriations committee has not passed on Mr Reagan's request for $20 million to be added to the Pentagon's construction funds for the current year to begin work on the production of binary chemicals weapons at the Pine Bluff arsenal in Arkansas, although this decision could be overturned in negotiations with the House of Representatives, which has already given its approval for the money.

The decision to add additional funds for the laser system is the result of growing pressure from the US military to respond to reports that the Soviet Union is already testing ways of using lasers to destroy missiles — and in the face of continuing scepticism from several quarters that space-based lasers can in fact be produced with the required degree of reliability and accuracy.

The proposed $50 million, which will include a detailed systems definition of a space-based laser weapons system, is a compromise between the $250 million being requested by some advocates of the system, and fiscal conservatives concerned about an unnecessarily large expansion of the arms budget, which the Senate approved at a record figure of $136,500 million for 1982.

Asked for comments on the original $250 million proposal, Dr John Foster of the Department of Defense's Defense Science Board said it was too soon to accelerate space-based laser development towards integrated space demonstration for any mission, and that "a push toward integrated on-orbit demonstrations would seriously endanger necessary [research and

0028-0836/81/210181-02$01.00

development] work''. However, according to Senator John Warner, the Defense Science Board has concluded that the potential utility of space-based lasers is significant and that an average of $50 million per year should be added to the programme.

The money for the binary chemical weapons facility which the Senate armed services appropriations subcommittee has left out of the supplemental request for the fiscal year 1981 would have been the first step towards renewed chemical weapons production which has been halted since 1968. The Administration is thought to be moving cautiously, aware of the public concern about the deployment of a new generation of chemical weapons in Europe. The main debate will probably take place later this year, when the details of next year's budget request, likely to contain substantially more funds for chemical weapons, are debated in Congress.

David Dickson

Soviet space

Romanian resources

Last Friday's launch of Soyuz-40, carrying into orbit Romanian cosmonaut–researcher Dumitru Prunariu for a week's stay aboard the Salyut-6 space station, completes the series of manned flights envisaged under the Comecon ''Interkosmos'' programme. Since the series was inaugurated with the flight of the Czech Vladimir Remek in March 1978, these ''international'' programmes have taken on a standard form, regularly including Earth-resources photography, astrophysics, biomedical studies and technological experiments. On each occasion there has been a smelting and crystallization experiment, using a Soviet ''Splav'' furnace and a capsule provided by the non-Soviet partner, with an appropriate patriotic code-name: ''Altai'' for Mongolia, ''Berolina'' for East Germany and in the latest case, ''Dacia'' for Romania.

Each joint programme, however, has also reflected the contribution made by the non-Soviet partner to the whole ''Interkosmos'' programme. In the case of Romania, the emphasis was initially on instrumentation. According to Dr Ioan Ursu, First Deputy Chairman of the Romanian National Council for Science and Technology, from the beginning of the programme in the late 1960s, Romania had specialized in the production of magnetometers, specially adapted spectrometers and radiation detectors. Romanian instruments, he said, had contributed considerably to the accumulation of data on the chemical composition of the upper atmosphere, variations of the Earth's magnetic field and particle physics. The current programme, according to Dr Eugen Mendescu, Secretary of the Romanian commission for space activity, will include the ''Astro'' experiment, which will attempt to capture super-heavy nuclei. Other innovations include an experiment on the behaviour of liquids in ''microgravitational'' conditions and the monitoring of dynamic disturbance of the space station itself, in order to determine possible effects on the experiments (the Polish ''Syrena'' crystallization experiments in 1978 seem to have run into trouble on this account).

What prospects the Soviet manned space programme now holds for international cooperation remains a mystery. Current plans for Soyuz–Salyut still include flights by a French and an Indian cosmonaut, and during last month's celebrations of the twentieth anniversary of Yurii Gagarin's flight it was suggested that Finland might also be invited to participate.

Future international cooperation may well have received a setback with the death last autumn of Academician Boris Petrov, head of the ''Interkosmos'' programme. Academician Petrov was a firm supporter of such international flights, unlike, for example, the President of the Soviet Academy of Sciences, Anatolii Aleksandrov, who on one occasion commented sourly that ''Solyuz is not a trolleybus''. Nevertheless, the options remain open. At a press conference in Berlin last month, cosmonauts Pavel Popovich and Vladimir Aksenov observed that a manned flight to Mars was a ''problem for the coming decades'' which could ''best be settled'' by means of international cooperation involving many countries.

Vera Rich

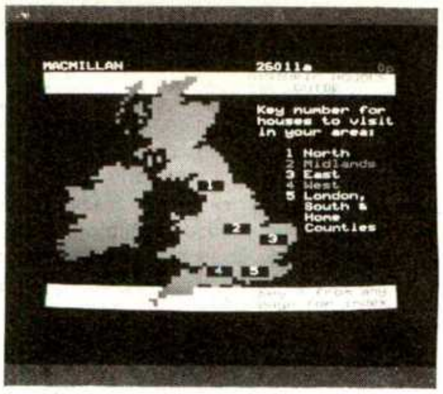
Viewdata systems

Battle joined

Although the British viewdata system, Prestel, was not expected to make a profit within the first eighteen months of operation, the results so far are disappointing. As yet only 1,288 people have bought the special television sets for domestic use. British Telecom's early hope that no home could do without its Prestel set seems to have been ill-founded. Sales to businesses, however, have been slightly brisker, although the 8,216 registered represent only a modest achievement.

British Telecom has, nevertheless, changed its marketing strategy, aiming at business rather than the domestic users. The National Economic Development Office hopes that business registrations will rise to 46,000 in the next eighteen months. The target for home registrations within that time is now 4,000.

The main lesson so far is that cost and the quality of information on the system are crucial to the home buyer. A custom-built Prestel set costs £400–500 more than an ordinary colour television set, although special adaptors costing about £200 are now coming onto the market. As well as ordinary telephone charges, a tariff for connection to a central computer is also levied, so the cost to home users who want only information on train timetables, local cinemas and so on can seem expensive. The recent decision to increase the computer connection charge for peak day-time use does not augur well.


MACMILLAN
26011a
Key number for houses to visit in your area:
1 North
2 Midlands
3 East
4 West
5 London, South & Home Counties
1
2
3
4
5

The quality of information at present stored on Prestel is variable. As yet nobody quite knows what type of information the domestic user is prepared to pay for, so two television supply companies are conducting a market survey in four provincial towns. Businesses, however, seem more prepared to pay for information, the greatest users so far being travel agents and financial companies.

One obstacle to further adoption by business is the increasing number of private viewdata systems, most of which are nevertheless compatible with Prestel. British Telecom regards these systems as complementary rather than competitive. It also hopes that the introduction next year of its Gateway system, whereby private databases can be linked to Prestel, will at least prevent British Telecom from being squeezed out of the market.

The teething problems of Prestel, the longest running public viewdata system in the world, will no doubt be watched with interest by competitors and potential buyers outside Britain. Although Prestel has been licensed to more telecommunications authorities than any of its rivals, the French Teletel and the Canadian Telidon may yet emerge on top. Arguments over which system should be recognized as a standard by the International Consultative Committee on Telephones and Telegraphs seem to have been resolved by all three being accepted. And the problem that no one system is compatible with either of the other two seems to have been resolved, at least for the French and British systems, by recent developments in software which will soon allow access by each system to the same computer.

Nevertheless, Telidon, which uses an alpha-numerical rather than an alpha-mosaic standard and is thus better at graphics, may have advantages over Prestel for some uses. What everyone wants to know is which system, if any, AT&T will choose in the United States. Some indication is expected to be given this week at Videotex '81 in Toronto.

Judy Redfearn

0028-0836/81/210182-01$01.00

Coal liquefaction

Trial by post

The future of the United Kingdom's only serious programme in coal liquefaction technology has been placed in doubt by a tense exchange of letters between the chairman of the National Coal Board (NCB), Sir Derek Ezra, and the Chief Scientist of the Department of Energy, Dr Tony Challis.

NCB has been waiting a year for a reply from the department about its proposal to build a 25-tonne-a-day pilot liquefaction plant at the Point of Ayr in North Wales, using two alternative technologies and costing £65 million to the government. These plans have become part of an agreement between the department, NCB and the National Union of Miners.

But, claims Dr Challis in his letter, on the assumption that oil conservation measures take full effect, and that sufficient oil-fired boilers are converted to coal-burning, "it seems likely that the United Kingdom can maintain net self-sufficiency in transport fuels and chemical feedstocks well into the next century", and thus that the United Kingdom does not need to develop an indigenous coal liquefaction technology, particularly when British coal is expensive (relative to American and Australian coal) and "bearing in mind the formidable level of international effort already applied in this field".

Dr Challis admits that NCB technology could be 30 per cent more energy-efficient than rival processes — after all, his own staff collaborated in the calculation — but says this is not enough to counter the 6-year lead-time in scale enjoyed by, say, the 200-tonne-a-day Exxon Donor Solvent plant at Baytown, United States. Dr Challis's letter, was written just after his return from a visit to this plant, so it seems Exxon have successfully "sold" him their technology.

He also argues that a 1-tonne-a-day plant followed by a 200-tonne-a-day plant would make better sense than the intermediate 25-tonne-a-day liquefier planned by NCB. Also, according to Dr Challis, who was a chemical engineer himself, NCB has "lacked adequate strength or perspective in the translation from chemistry to chemical engineering", and the translation has been done in too much of a hurry.

Sir Derek Ezra, in reply to Dr Challis, chides him for the department's late change of heart — and expresses surprise at Challis's optimistic view of Britain's oil supply, "more optimistic than the official view of the Department of Energy". And it is not at all clear, says Sir Derek, that countries with cheap surplus coal would be prepared to supply it cheaply in the future, when coal is in greater demand — so the cost of our own coal is not relevant.

Sir Derek believes that liquefaction technology will be needed in the United Kingdom in the 1990s, and that NCB technology presents "a sufficient advance over the technologies being developed in the United States and Germany to merit support".

The American technology — and a similar plant under construction at Bottrop in West Germany — is, in fact, based largely on the wartime German technology (also employed at SASOL in South Africa) which splits the coal into monocarbon compounds, resynthesizing them to make light liquids. The resulting liquids possess only 30 per cent or so of the chemical energy of the original coal, whereas the NCB process which takes coal directly to light liquids, can promise 60–70 per cent efficiency.

British Petroleum is prepared to back the project, while NCB itself will contribute £18 million, and the project would receive about £12 million from the energy research fund of the European Community, if only the Department of Energy would give it their approval. Moreover, potential buyers of the technology in Canada and Australia will only show serious interest when a pilot plant is built.

However, despite Dr Challis's views, the department has not yet made a formal decision about the plant. The next step should be a meeting shortly between the minister for coal John Moore, Sir Derek Ezra, Dr Challis and the research director of NCB, Dr Joe Gibson. If they can reach agreement, their recommendation will then be forwarded to the Secretary of State for Energy for a decision.

Meanwhile, in the United States President Reagan has appointed an Oklahoma businessman, Mr Edward E. Noble, to be chairman of the Synthetic Fuels Corporation which will oversee synfuel development in the United States. Mr Noble is reported to be uneasy at the large numbers of US technologies aimed at producing gas, and uncertain that the mandated targets of 0.5 million barrels a day (oil equivalent) of synfuel by 1987 and 2 million by 1992 can be reached.

In London, the director of the International Energy Agency said that given uncertainties over nuclear and renewable sources of energy, coal was poised to become the "single most important energy contribution to the future of modern society".

Robert Walgate

From rank weeds to riches

New Delhi

Workers in India have found a way of converting water hyacinth (*Eichornia crassipes*) into paper, and in so doing may have solved one of the Third World's major problems. This free-floating and fast-growing weed is not only a menace to navigation and hydroelectric power generation in affected countries, but also offers a breeding ground for pests and insects. So far, all efforts have been directed at eradication of the weed. Now, the Regional Research Laboratory (RRL) in Hyderabad, controlled by the Council of Scientific and Industrial Research, is making writing paper, posters, even invitation cards, using the leaves of water hyacinth collected from the city's lakes. Because the raw material is virtually free the paper is cheap and may save some of the trees cut down to feed the paper industry.

Water hyacinth — paper power?

The technology for converting the weed into paper was developed through an international project initiated in 1978 by the London-based Commonwealth Science Council. The council, which identified water hyacinth as a major problem in the Third World, asked India to coordinate a project on the control and utilization of the weed and helped the collaborating countries to obtain a grant of $200,000 from the United Nations Development Programme. While Sri Lanka, Bangladesh and Malaysia worked on biological control and biogas production, India chose to investigate the use of the weed in paper making. And, when scientists from the collaborating countries recently visited the RRL pilot plant, they were so impressed that they now plan to set up similar plants in their own countries.

The emergence of water hyacinth as a source of raw material for the paper industry should have great impact in India, where at least four million hectares of water surface are covered with the weed. The average yield of water hyacinth is 50 tonnes per hectare, giving India 200 million tonnes of non-forest raw material for paper. If even half of this is used, and assuming a 10 per cent conversion efficiency, the country can provide 10 million tonnes of paper from the weed. The new paper-making factories can be set up wherever water hyacinth grows and the process is such that the paper can be hand-produced by unskilled workers.

K.S. Jayaraman

0028-0836/81/210183-01$01.00

The responsibility of scientists

from Andrei Sakharov

Because of the international nature of our profession, scientists form the one real worldwide community which exists today. There is no doubt about this with respect to the substance of science: Schrodinger's equation and the formula $e = mc^2$ are equally valid on all continents. But the integration of the scientific community has inevitably progressed beyond narrow professional interests and now embraces a broad range of universal issues, including ethical questions. And I believe this trend should and will continue.

Scientists, engineers and other specialists derive from their professional knowledge and the advantages of their occupations a broad and deep understanding of the potential benefits — but also the risks — entailed in the application of science and technology. They also develop an awareness of the positive and negative tendencies of progress generally, and its possible consequences.

Colossal opportunities exist for the application of recent advances in physics, chemistry and biochemistry; technology and engineering; computer science; medicine and genetics; physiology and hygiene; microbiology (including industrial microbiology); industrial and agricultural management techniques; psychology; and other exact and social sciences. And we can anticipate more achievements to come. We all share the responsibility to work for the full realization of the results of scientific research in a world where most people's lives have become more difficult, where so many are threatened by hunger, premature illness and untimely death.

The dangers of progress

But scientists and scholars cannot fail to think about the dangers stemming from uncontrolled progress, from unregulated industrial development, and especially from military applications of scientific achievements. There has been public discussion of topics related to scientific progress: nuclear power; the population explosion; genetic engineering; regulation of industry to protect the environment; protection of air quality, of flora and fauna, and of rivers, lakes, sea and oceans; the impact of mass media. Unfortunately, despite the urgent and serious nature of the issues at stake, such discussions are often uninformed, prejudiced or politicized, and sometimes simply dishonest. Experts, therefore, are under an obligation to subject these problems to unbiased and searching examination, making all socially significant information available to the public in direct, first-hand form, and not just in filtered versions. The discussion of nuclear power, a subject of prime importance, is an instructive example. I have expressed elsewhere my opinion that the dangers of nuclear power have been exaggerated in the West, and that such distortion is harmful.

With some important exceptions (primarily affecting totalitarian countries), scientists are not only better informed than the average person, but also strive for and enjoy more independence and freedom. Freedom, however, always entails responsibility. Scientists and other experts already influence or have the capacity to influence public opinion and their governments. (That influence should not be exaggerated, but it is substantial.) My view of the situation of scientists in the contemporary world has convinced me that they have special professional and social responsibilities. It is often difficult to separate one from the other — the communication of information, the popularization of scientific knowledge, and the publication of endorsements or warnings are examples of activities with both professional and social aspects.

Similar complications arise when scientists become involved in questions of disarmament: in developing strategy for or participating in international negotiations; in advancing proposals or issuing appeals to governments or to the public; and in alerting them to dangers. Disarmament is a separate, critically important issue which requires a profound, thorough and scientifically daring approach. I realize that more detailed treatment is needed, but now I will simply outline a few ideas. I consider disarmament necessary and possible only on the basis of strategic parity. Additional agreements covering all kinds of weapons of mass destruction are needed. After strategic parity in conventional arms has been achieved, a parity which takes account of all the political, psychological and geographical factors involved, and if totalitarian expansion is brought to an end, then agreements should be reached first prohibiting the use of nuclear weapons, and later, banning such weapons.

Defence of human rights

Another subject which is closely connected to questions of peace, trust and understanding among countries is the international defence of human rights. Freedom of opinion, freedom to exchange information and freedom of movement are necessary for true accountability of the authorities which in turn prevents abuses of power in domestic and international matters. I believe that such accountability would make impossible tragic mistakes like the Soviet invasion of Afghanistan and would inhibit manifestations of an expansionist foreign policy and acts of internal repression.

The unrestricted sale of newspapers, magazines and books published abroad would be a major step towards effective freedom of information in totalitarian countries. Perhaps even more significant would be the abolition of censorship which should concern first of all the scientists and intelligentsia of totalitarian countries. It is important to demand a halt to jamming of foreign broadcasts which deprives millions of access to the uncensored information needed to form an independent judgment of events. (Jamming was resumed in the USSR in August 1980 after a seven-year interval.)

I am convinced that support of Amnesty International's call for a general, worldwide amnesty for prisoners of conscience is of special importance. The political amnesties proclaimed by a number of countries in recent years have helped to improve the atmosphere. An amnesty for prisoners of conscience in the USSR, in Eastern Europe, and in all other countries where political prisoners or prisoners of conscience are detained would not only be of major humanitarian significance but could also enhance international confidence and security.

The worldwide character of the scientific community assumes particular importance when dealing with such problems. By its international defence of persecuted scientists and of all persons whose rights have been violated, the scientific community confirms its international mandate which is so essential for successful scientific work and for service to society.

Western scientists are familiar with the names of many Soviet colleagues who have been subjected to unlawful repressions. (I shall confine my discussion to the Soviet Union since I am better informed about it, but serious human rights violations occur in other countries including Eastern European countries.) The individuals I mention have neither advocated nor used violence since they consider publicity the only acceptable, effective and nonpernicious way of defending human rights. Thus, they are all prisoners of conscience as defined by Amnesty International. Their stories have much else in common. Their trials were conducted in flagrant violation of statutory procedures and in defiance of elementary common sense. My friend Sergei Kovalev was convicted in 1975 in the absence of the defendant and counsel, that is, with no possibility whatsoever of defence. He was sentenced to seven years in a labour camp and three years of internal exile for anti-Soviet agitation and propaganda allegedly contained in the samizdat news magazine *A Chronicle of Current Events*, but there was no examination of the substance of the charge.

Comparable breaches of law marked the trials of Yury Orlov, the founder of the Moscow Helsinki Group, and of other members of the Helsinki Groups and associated committees: Victor Nekipelov, Leonard Ternovsky, Mykola Rudenko, Alexander Podrabinek (and his brother Kirill), Gleb Yakunin, Vladimir Slepak, Malva Landa, Robert Nazarian, Eduard

0028-0836/81/210184-02$01.00

Arutyunian, Vyacheslav Bakhmin, Oles Berdnik, Oksana Meshko, Mykola Matusevich and his wife, and Miroslav Marinovich. Tatiana Osipova, Irina Grivnina and Felix Serebrov have been imprisoned pending trial. Yury Orlov's lawyer missed part of the trial proceedings when he was locked up forcibly in chambers adjoining the courtroom. Orlov's wife was frisked in a crude way and her clothing ripped during a search for written notes or a tape recorder, all from fear that the court's grotesque secrets might be revealed.

Mistreatment

In the labour camps, prisoners of conscience suffer cruel treatment: arbitrary confinement in punishment cells; torture by cold and hunger; infrequent family visits subject to capricious cancellation; and similar restrictions on correspondence.

E. Yantelevich

Andrei Sakharov at Gorky

They share all the rigours of the Soviet penal regimen for common criminals while suffering the added strain of pressure to "embark on the path of reform", that is, to renounce their beliefs. I would like to remind you that not once has any international organization, such as the Red Cross or a lawyers' association, been able to visit Soviet labour camps.

Political prisoners are often rearrested, and monstrous sentences imposed. Ornithologist Mart Niklus, poet Vasyl Stus, physics teacher Oleksei Tikhy, lawyer Levko Lukyanenko, philologist Viktoras Petkus and Balys Gajauskas have all received sentences of ten years in labour camp and five years internal exile as recidivists. A new trial is expected for Paruir Airikian who is still in labour camp. Within the past few days I have been shocked by the fifth (!) arrest of my friend Anatoly Marchenko, a worker and author of two talented and important books: *My Testimony* and *From Tarusa to Siberia*. Imprisoned religious believers include Rostislav Galetsky, Bishop Nikolai Goretoi, Alexander Ogorodnikov, and Boris Perchatkin. Imprisoned workers include Yury Grimm and Mikhail Kukobaka. Alexei Murzhenko and Yury Fedorov are still imprisoned. I shall name only a few scientists deprived of their freedom; many others could be added to the list: Anatoly Shcharansky, the young computer scientist now famous around the world; mathematicians Tatiana Velikanova, Alexander Lavut, Alexander Bolonkin and Vazif Meilanov; computer scientist Victor Brailovsky; economist Ida Nudel; engineers Reshat Dzhemilev and Antanas Terleckas; physicists Rolan Kadiyev, Iosif Zisels and Iosif Dyadkin; chemists Valery Abramkin and Juri Kukk; philologists Igor Ogurtsov and Mustafa Dzhemilev; and Vladimir Balakhonov.

A common violation of human rights, and one which especially affects scientists, is denial of permission to emigrate. The names of many "refusniks" are known to the West.

I was banished without a trial to Gorky more than a year ago and placed under a regimen of almost total isolation. A few days ago the KGB stole my manuscripts and notebooks which contained extracts from scientific books and journals. This is a new attempt to deprive me of any opportunity for intellectual activity, even in my solitude, and to rob me of my memory. For more than three years Elizaveta Alexeyeva, my son's fiancée, has been arbitrarily prevented from leaving the Soviet Union. I have mentioned my own situation because of the absence of any legal basis for the actions taken and because the detention of Elizaveta is undisguised blackmail directed against me. She is a hostage of the state.

Fight repression

I appeal to scientists everywhere to defend those who have been repressed. I believe that in order to protect innocent persons it is permissible and, in many cases, necessary to adopt extraordinary measures such as an interruption of scientific contacts or other types of boycott. I urge the use, as well, of all the possibilities of publicity and of diplomacy. In addressing the Soviet leaders, it is important to take into account that they do not know about — and probably do not want to know about — most letters and appeals directed to them. Therefore, personal interventions by Western officials who meet with their Soviet counterparts have particular significance. Western scientists should use their influence to press for such interventions.

I hope that carefully thought out and organized actions in defence of victims of repression will ease their lot and add strength, authority and energy to the international scientific community.

I have entitled this letter "The Responsibility of Scientists". Tatiana Velikanova, Yury Orlov, Sergei Kovalev and many others have decided this question for themselves by taking the path of active, self-sacrificing struggle for human rights and for an open society. Their sacrifices are enormous, but they are not in vain. These individuals are improving the ethical image of our world.

Many of their colleagues who live in totalitarian countries but who have not found within themselves the strength for such struggle, do try to fulfil honestly their professional responsibilities. It is, in fact, essential to work at one's profession. But has not the time come for those scientists, who often exhibit their perception and nonconformity when with close friends, to demonstrate their sense of responsibility in some fashion which has more social significance, and to take a more public stand, at least on issues such as the defence of their persecuted colleagues and control over the faithful execution of domestic laws and the performance of international obligations? Every true scientist should undoubtedly muster sufficient courage and integrity to resist the temptation and the habit of conformity. Unfortunately, we are familiar with too many counterexamples in the Soviet Union, sometimes using the excuse of protecting one's laboratory or institute (usually just a pretext), sometimes for the sake of one's career, sometimes for the sake of foreign travel (a major lure in a closed country such as ours). And was it not shameful for Yury Orlov's colleagues to expel him secretly from the Armenian Academy of Sciences while other colleagues in the USSR Academy of Sciences shut their eyes to the expulsion and also to his physical condition? (He is close to death.) Many active and passive accomplices in such affairs may themselves someday attract the growing appetite of Moloch. Nothing good can come of this. Better to avert it.

Western scientists face no threat of prison or labour camp for public stands; they cannot be bribed by an offer of foreign travel to forsake such activity. But this in no way diminishes their responsibility. Some Western intellectuals warn against social involvement as a form of politics. But I am not speaking about a struggle for power — it is not politics. It is a struggle to preserve peace and those ethical values which have been developed as our civilization evolved. By their example and by their fate, prisoners of conscience affirm that the defence of justice, the international defence of individual victims of violence, the defence of mankind's lasting interests are the responsibility of every scientist.

Gorky 24 March 1981

P.S. After this letter was written, I received word of the tragic death of Juri Kukk in a labour camp. Tatiana Osipova was sentenced to five years in a labour camp and five years internal exile on 2 April.

© 1981 Andrei Sakharov

This essay was written by Andrei Sakharov for delivery to the International Conference in Honor of Andrei Sakharov sponsored by The New York Academy of Sciences, the American Institute of Physics and the American Physical Society held on 1–2 May 1981, at Rockefeller University, New York. (English translation courtesy of Khronika Press.)

CORRESPONDENCE

Creationists persist

Sir — No doubt the opponents of evolution are delighted with the fireworks of disagreement that are now being displayed on this subject by biologists who write to *Nature*. Indeed, the critics of Copernicus took heart 400 years ago because scientists could not agree among themselves. However, it is not likely that the scientific process of argument betokens a swing to creationism.

During the Sacramento trial, 2–6 March, the creationists roared as gently as any sucking dove. They said, in the courtroom, that they were not seeking equal time for the teaching of creationism with evolution. In the corridors, talking to reporters, it was a different story. Their veteran leader, Mrs Nell Segraves, said: "We want 50 per cent of the tax dollar used for education to our point of view. We have a lot to undo. Creation/evolution is only the beginning."

The head of the Institute for Creation Research, San Diego, Henry Morris, PhD (Hydraulic Engineering), does not mince matters when he is writing for an audience of believers. He has said that "It is high time that Christians face the fact that the so-called geologic ages are essentially synonymous with the evolutionary theory of origins. The latter, in turn, is, at its ultimate roots, the anti-God conspiracy of Satan himself." He also says the fractures and scars of the Moon and Mars reflect a fight between good and bad angels.

However, the argument used for persuading school boards and legislatures is that "creation science" is a credible variant of conventional science, and deserves equal time. A creationist textbook has anticipated and countered Ashley Montagu's plea (*Nature* 23 April, p.623) that "perhaps the most remarkable and obvious [example of evolution in process] is the well-known case of industrial melanism in moths". Not so, say the creationists, "The moths are just the same as they were before the Industrial Revolution, and they are not becoming anything different . . . Darwin devoted his life in attempting to make it seem reasonable that all forms of life on Earth evolved from one or a few simple forms; this is a very different matter from birds selecting dark or light moths according to the type they can see more readily on a certain background". This argument is strengthened by recent observations that the black morphs are now disappearing because soot is disappearing from trees in the formerly smoky and unkind Midlands.

Creationism is on the march in Oregon, Florida, and Michigan. Governor Frank White signed the Arkansas Bill on 19 March requiring the "balanced treatment" of "creation-science" and evolution in public schools, and admitted afterwards that he had not read all of it. However, on 29 March he told a church congregation that creation-science "is a theory supported by documentation, and evolution-science is the same".

The "equal time for nonsense" position is not one that favours science; science is governed by merit, not fairness. As pointed out by Weber, the position that has the best evidence, has withstood prolonged criticism, and has the mass support from knowledgeable workers in the field, is the position that should be given the emphasis in education. In science, one theory is *not* just as good as another.

Thomas H. Jukes

University of California,
Berkeley, USA

Theory upon theory

Sir — The author of the article "How true is the theory of evolution" (*Nature* 12 March, p.75) seemed to misapprehend Popper's refutation principle. He referred not only in the title to the "truth" of a theory but in the article also, where he compounded the error by writing about "proving" theories. Truth and Proof are precisely what Popper's refutation principle does not establish.

It is also unnecessarily narrow to apply Popper's demarcation principle as the exclusive arbitrator in a dispute over the scientific content of a theory. There is no known test or property that demonstrates unequivocally that a theory is scientific or otherwise (see, for example, *The Structure of Scientific Theory*, ed. F. Suppe, 1977). Popper's demarcation principle is not exceptional in this respect; it is probably an analysis most suitably used to explore the scientific content of the theories of physics rather than evolutionary theories, the latter have a historical component the former lack.

Historical theory is often formally different from scientific theory as defined by Popper, as the former can sometimes be confirmed by a finite number of observations. They are confirmed or refuted when all possible observations have been made. In this type of theory, refutation has no logical priority over confirmation — the state of affairs Popper claims for the theories he labels as scientific.

An example of the confirmation/refutation symmetry in historical theory may be found in the theory of evolution, where the gaps in the fossil record are explained sometimes as undiscovered transitional forms. This is theoretically a confirmable historical theory, though in practice the examination of every piece of the Earth's sedimentary rock would be an impossible task. In reality, this theory becomes irrefutable and therefore not scientific according to Popper, unless we interpret his "falsifiable in principle" to mean theoretically testable even if not testable in practice. Popper seems to interpret it thus, as in the *New Scientist* (21 August 1980) he claims historical theories are scientific, but this seems to upset his claim for the priority of refutability over confirmation as a property of scientific theories. On the other hand, should the interpretation be that the explanation of the gaps as transitional forms was in practice irrefutable, it is difficult to accept that finding transitional forms that fit the gaps is in no way a scientific activity.

I would also like to draw to the attention of the author of the article in *Nature* a conclusion of Feyerabend, for example, who states that "relativity theory does not modify Newtonian mechanics, it displaces the theory altogether". The displacement of one major theory by another, according to him, occurs because the two theories represent mutually exclusive universes of discourse and therefore do not have a common language. In this instance, because the two theories embrace incompatible concepts of motion, Newtonian mechanics and the theory of relativity therefore on Feyerabend's analysis do not have an analogous relationship to Lamarckism and neo-Darwinism. The author's conclusion that Lamarckism would modify neo-Darwinism was correct, but the analogy chosen was probably inappropriate.

Barrie Pearson

Winfrith, Dorchester, UK

Use the library

Sir — The recent letter from Andrew Brooks[1] indicates problems which are encountered by everyone engaged in research. As a Librarian/Information Officer in a small research institute, I often find that the information aspect of research work is approached the wrong way round. The laboratory work is done and a paper almost written before the library is consulted for relevant references. Senior research staff can be particularly guilty, relying strongly on the "old boy" network.

The information explosion of the past 15 years has greatly complicated the task of obtaining relevant information and sophisticated systems have been developed to make the job easier. *Chemical Abstracts* is perhaps the best known information retrieval tool in the sciences. This includes about half a million abstracts per year and anyone not knowing how to use it will have difficulty in retrieving any relevant information. Computerized data bases abound and a directory has been published[2]. *Chemical Abstracts* can also be searched on-line[3].

Training in how to obtain information from published sources is surely now essential. I feel it is time for a compulsory, assessed course in information retrieval for undergraduates to be introduced along with other core subject courses.

Gill Croft

International Tin Research Institute,
Greenford, Mddx, UK

1. Brooks, A. *Nature* **291**, 7 (1981).
2. Hall, J.L. *On-line Bibliographic Data Bases* (Aslib, London, 1981).
3. Barker, F.H. *Aslib Proc.* **31**, 538 (1979).

Sir — Several months to carry out a literature search! Andrew Brooks' letter (*Nature* 7 May, p.7) amazes me and seems to reveal a lack of initiative by him and by his library. It is part of the job of an academic librarian to explain information handling techniques, whether informally or via a specialist course. In addition, most of us have access to computerized information sources and use them routinely to undertake literature searching for teaching and research staff. Surely a first task for any new researcher (or, indeed, lecturer) is to find out what services the library offers, just as he will investigate available computing facilities. Did Mr Brooks not do this? Conversely, though there are obviously difficulties in institutions with large numbers of researchers, librarians should seek out and make contact with new staff to explain their services. Nevertheless, the SRC might still usefully insist on a course of research methods, including techniques of information retrieval, as an initial part of any PhD programme.

I.R. Winship

Polytechnic Library,
Newcastle-upon-Tyne, UK

Genes of lymphocytes II:
T-cell receptors, idiotypes and the MHC

from Miranda Robertson

OF the inexhaustible enigmas of the immune system, the focal one is certainly the T-cell receptor. Much recent work on the DNA of cloned T cells has been inspired by the hope that T cells, like B cells, recognize antigen by means of binding sites encoded by immunoglobulin variable-region genes. It is well established that T cells do not bear surface immunoglobulin, however, and the evidence that they express V genes is indirect. It is also central to the principal unifying theory of immune regulation.

The theory, originally formulated by Jerne for B cells alone[1], is based on the notion of an interacting network of antibodies which recognize antigenic determinants, known as idiotypes, on the antigen-binding sites of other antibodies. Binding to the idiotype by an anti-idiotype antibody may either induce or suppress the proliferation of the cell bearing it. But much of the immune response and its regulation depend on the activities of T cells: cytotoxic T cells, which kill infected cells; helper T cells, which induce the proliferation of cytotoxic T cells and of some B cells; and suppressor T cells, which suppress all these responses, probably in most cases by acting on the T helper cells that promote them.

One of the most important discoveries of the past decade, therefore, was the demonstration that T-cell help in antibody production could be induced or inhibited by an injection of anti-idiotype antibody[2]. This suggested that T cells bear the same idiotypes, and thus by implication the same antigen-binding sites, as B cells; and made it possible to see how they could participate in the idiotype–anti-idiotype regulatory network proposed by Jerne.

The T-cell receptor

The idea of a T-cell regulatory network based on idiotypes has been enthusiastically embraced by most cellular immunologists, although not all attempts to control immune responses with anti-idiotypic antisera have been uniformly successful. In the hope of clarifying both the cellular interactions and the molecular signals underlying such regulation, several laboratories have isolated antigen-specific T suppressor and T helper cells which carry idiotype-bearing factors on their surfaces. These factors are shed into the medium of T cells *in vitro,* or extracted from the cells, and can be shown to substitute for them in antigen-specific suppression or help.

The biochemical characterization of T-cell factors has, however, been slow, and is now being rapidly overtaken by the investigation of cloned T-cell lines at the DNA level. The rationale for these investigations is that if T cells are expressing receptors encoded by V-region genes, the V-region DNA of T cells should be rearranged as it is in B cells (see ref. 3 for details of the genes and their rearrangements). The idiotypes expressed by T cells seem to be encoded by heavy-chain V genes of immunoglobulin, so the investigators have used heavy-chain DNA probes to look at T-cell DNA. So far, the results have been negative.

Kurosawa *et al.*[4], using a J-region probe, have failed to find any functional rearrangement in a cytotoxic T-cell line. This is perhaps not entirely surprising, since cytotoxic T cells do not seem to express immunoglobulin idiotypes (though this in itself is a little unsettling for a network advocate). But at Salt Lake City*, T. Honjo (Osaka University) also reported, on the basis of preliminary data, failure to find rearrangements in T helper-cell lines.

What might these failures mean? One possibility is that the idiotypes expressed on T cells have nothing to do with the V-region genes of immunoglobulins. I. Weissman (Stanford University), playing devil's advocate, pointed out that the T15 idiotype characteristic of the anti-phosphorylcholine antibodies of certain mouse strains is recognized by a monoclonal antibody against Thy-1 antigen — a kind of cross-reaction that has been pointed out before in connection with other cell-surface molecules.

The alternative possibility, which depends on the fact that so far only J-gene probes have been used, is that in T cells, V_H genes recombine not with J_H genes but with an unknown set of T-cell J genes (J_T genes) elsewhere on the DNA. A report from F. Owen (Tufts University) on a genetically defined, heavy chain-linked T-cell marker mapping downstream of the Ig constant-region genes led to speculation that this might be the site of clusters of the putative J_T and C_T genes.

In any case, one might not expect to find precisely the same genes encoding the antigen-binding regions in T- and B-cell receptors, since the T-cell receptor repertoire is known to be more limited than that of the B cell. One of its most profound limitations is its inability to recognize antigen except when the antigen is associated with a 'self' major histocompatibility complex (MHC) molecule. The mechanism and ontogeny of this so-called MHC restriction are among the most disputed issues of modern immunology, the central questions being whether T cells see both foreign antigen and self-MHC antigen with the same or separate receptors; and whether the receptor repertoire is selected in the thymus — since selected it must be, in order that self MHC be tolerated when free of foreign antigen and attacked when associated with it.

An additional complication for immunologists working on idiotype-bearing T cells and their factors, therefore, is that they must show the appropriate MHC restriction. There are, in fact, several antigen-specific, idiotype-bearing T-cell factors; J. Kapp and her collaborators (Jewish Hospital of St Louis) have recently purified a 25,000-MW suppressor factor specific for GAT and bearing a mouse MHC I-J haplotype, that is particularly interesting because it acts in sub-nanogram quantities *in vivo.*

One of the questions constantly raised about T-cell factors in particular and anti-idiotypic regulation in general is whether they have any physiological relevance. Even the original Eichmann–Rajewsky experiments[2] are open to criticism on the grounds that they were done with xenogeneic anti-idiotype antisera. In their most recent experiments, Rajewsky and his colleagues used nanogram quantities of syngeneic anti-idiotypic antibodies to

---

Miranda Robertson is an associate editor of Nature.

*The ICN/UCLA Symposium on 'Immunoglobulin Idiotypes and their Expression' was held in Salt Lake City, Utah, on February 8–15, 1981. A report on contributions dealing with immunoglobulin genes has already appeared[3].

0028-0836/81/210187-02$01.00

enhance or suppress — depending on timing of administration and dosage of antibody — the responses of mice to hydroxynitrophenol (NP)[4]. Their very latest achievement is to have enhanced the idiotype-positive component of an antibody response to NP by injection of 10 μg of idiotype-bearing antibodies in advance of immunization with NP.

The major histocompatibility complex

If the T-cell receptor remains elusive, the MHC is about to yield up its mysteries to the inexorable advance of the DNA-cloning army. So far, however, the process has only just started, the sequencing has barely begun, and the data largely confirm the existing biochemistry — though there may be some surprises for the geneticists.

MHC antigens, known as HLA antigens in man and H-2 in mouse, are composed of two chains, the smaller of which, β_2 microglobulin, is relatively invariant. The structure of the protein chain shows considerable homology to the constant region of immunoglobulin, but a preliminary analysis of the human genomic DNA (J. Parnes, US National Institutes of Health) suggests that the distribution of introns may be more like that of the variable region.

The larger, α chain, which is highly polymorphic, has only more recently been recognized as having structural similarities to immunoglobulin[5]. Comparing amino acid sequences, J. Strominger (Harvard University) finds 35 per cent homology between the third domain of the α chain of HLA-A and the fourth domain of the immunoglobulin constant region. L. Hood (California Institute of Technology), who has used an HLA probe to identify the H-2 genes, has sequenced the cDNA coding for the third α domain and finds 50 per cent homology with the DNA encoding the fourth constant-region domain of mouse IgM. Moreover, he infers from analysis of the 'wobble' positions of the codons, which have 45 per cent homology instead of the 25 per cent that would be expected by chance, that the homology is due to divergent rather than convergent evolution.

The most startling news came from the California Institute of Technology where, according to E. Moore, densitometric analysis of genomic bands from the H-2 complex of the mouse suggest there may be 15 or so related genes[6] rather than the four or five predicted by the genetics. Similar data have since been reported by Cami *et al.*[7], and probably signify the existence of a multi-gene family possibly encoding lymphocyte differentiation antigens.

This raises the question, posed by L. Hood (California Institute of Technology), of how to identify the members of a multi-gene family without knowing what proteins they encode — a problem to which J. Sim (Columbia University) supplied the solution, at least for membrane proteins. She described a system for using cloned DNA to transform fibroblasts and for assaying the fibroblasts for new surface antigens with cytotoxic lymphocytes, or with antibodies — a technique she and her colleagues have used to identify a genomic clone containing an HLA antigen-encoding sequence.

Nothing in immunology is beyond the wit of man (or woman). T-cell receptor next. □

---

1. Jerne, N.K. *Scient. Am.* 52 (1973).
2. Eichmann, K. & Rajewsky, K. *Eur. J. Immun.* **5**, 661 (1975).
3. Robertson, M. *Nature, News and Views* **290**, 625 (1981).
4. Reth, M., Kelsoe, G. & Rajewsky, K. *Nature* **290**, 257 (1981).
5. Steinmetz, M. *et al. Cell* **24**, 125 (1981).
6. Cami, B., Brégégère, F., Abastado, J.P. & Kourilsky, P. *Nature* (in the press).

Blood lead concentrations reconsidered

from Philippe Grandjean

Two recent studies of blood lead concentrations have provided evidence that, compared with the usual average of 10–20 μg dl^{-1} found in Western Europe and North America, levels in remote societies are much lower. Poole and colleagues[1] found a mean blood lead concentration of 5 μg dl^{-1} in 100 children from an unpolluted area in Papua New Guinea and Piomelli and coworkers[2] established an average blood level of about 3 μg dl^{-1} in 103 Nepalese children and adults from the foothills of the Himalayas. The two investigations were carefully performed in order to avoid the contamination that has marred many lead assays in the past[3]. The results are supported by an earlier study of Venezuelan Indians[4] which also indicated low blood lead concentration — below 1 μg dl^{-1} — although the analytical techniques were perhaps subject to some doubt.

Atmospheric lead pollution spreads throughout the world so even population groups living in remote areas may be exposed to lead levels which are higher than natural, pre-industrial levels. One way in which natural levels can be assessed is by the analysis of prehistoric skeletal remains. Such samples, if uncontaminated, can be used to give accurate estimates, since lead accumulates in calcified tissues. Determination of lead in temporal bones and secondary dentine found in the sands of the Sahara has enabled lead levels in Sudanese Nubia to be assessed for the period from BC 3300 to AD 750 (ref.5). The results show that about 5,000 years ago, lead exposure at the Nile banks in Nubia was about one per cent of current exposures in Western Europe and North America. Later, during the Egyptian occupation and with the introduction of metallurgical techniques, tissue lead levels in Nubia increased considerably.

Further evidence comes from Ericson *et al.*[6] who examined calcified tissues of a small number of Peruvian mummies. Using sophisticated analyses, they estimated that present-day lead exposures may have increased by about 500-fold above pre-metallurgical levels in Peru.

A report from the US National Academy of Sciences (NAS)[7] also carefully evaluates the natural occurrence of lead in the environment. Without doubt, lead was never uniformly distributed in the biosphere, and even in pre-industrial times a range of natural exposures would be expected. The NAS committee concludes that current exposure levels in industrialized countries are about 10- to 1,000-fold above natural concentrations.

What is the significance of these figures for human health? Unfortunately, no clear answer is available. Although current background lead levels appear well above those which early man experienced, there is no definite evidence for the threshold at which lead begins to have a toxic effect. Most studies have examined effects at, say, concentrations of lead 1,000–10,000 times the natural level and compared them to a control group with a lead burden 100-fold greater than that. However, toxic effects of lead are now being found at increasingly low levels and the most recent evidence[8] suggests that they may be close to, or perhaps even below, current background levels.

Until proper studies of low-level lead toxicity, incorporating 'unpolluted' control groups, can be conducted, the recommendation of the NAS committee that "a serious effort should be made to reduce the baseline level of exposure to lead for general population of the United States" should be carefully considered in every industrialized country. □

Philippe Grandjean is in the Department of Occupational Medicine, Danish National Institute of Occupational Health.

1. Poole *et al. Sci. Total Envir.* **15**, 17 (1980).
2. Piomelli *et al. Science* **210**, 1135 (1980).
3. Seattle *et al. Science* **207**, 1167 (1980).
4. Hecker *et al. Arch. envir. Hlth* **29**, 181 (1974).
5. Granjean *et al. J. envir. Path. Tox.* **2** (1979).
6. Ericson *et al. New Engl. J. Med.* **300**, 946 (1979).
7. *Lead in the Human Environment* (US National Academy of Sciences, Washington DC, 1980).
8. Needleman *Low Level Lead Exposure* (Raven, New York, 1980).

0028-0836/81/210188-01$01.00

Continental crust compromise

Pure Chemistry Library
SCIENCE COLLEGE
92, Upper Circular Road
CALCUTTA

from Keith Bell

THE evolution of the continental crust has been a subject of hot debate. Numerous models have been proposed (see, for example, refs 1–4), but most can be roughly placed into one of two groups. Proponents of the first group believe that most continental crust formed fairly early in Earth history, perhaps within the first 500 Myr or so, and that its subsequent reworking gave rise to a variety of granitic products. The alternative proposal argues for a more gradual evolution with additions to the crust being made at intervals throughout its history, either from island arcs at an accreting margin ('lateral accretion') or directly from below ('vertical accretion'). Implicit in some of these models is either an upper mantle that changed continuously or one that had its major differentiation shortly after the formation of the Earth. Also raised in this whole discussion is the origin of granites: are they extracted directly from the mantle? Do they simply represent reworked, older continental crust or are they produced by the mixing of both mantle and crustal components? Recent Sm–Nd work by DePaolo, reported on p.193 of this issue of *Nature*, answers at least some of these questions.

Gast's classic paper[5] was followed by further attempts to place constraints on the evolution of the crust–mantle system by isotopic means. One obvious approach was to use initial ^{87}Sr/^{86}Sr ratios but this produced surprising results: a large proportion of granitoid rocks were shown to have fairly low and variable initial ^{87}Sr/^{86}Sr ratios (<0.706) — a feature which, in its simplest interpretation, favoured a direct mantle origin. Low initial Sr ratios are, however, open to several interpretations including derivation of granites from a source with a similar, low Rb/Sr ratio to that of the mantle (lower crust ?); the mixing of continental material with 'primitive' Sr contained either in the mantle, lower crust or within a fluid phase; or derivation from a terrane that somehow had been swept clean of radiogenic Sr (granulite terranes ?). Isotopic data from one radiometric system were obviously inadequate to resolve many of the problems of crustal genesis, but the use of initial ^{87}Sr/^{86}Sr ratios in conjunction with common Pb data still hinted at a direct mantle origin for most granites[6].

A new and powerful tool for evaluating the evolution of the upper mantle and the origin of the crust has now been provided by the Sm–Nd method[7]. Because of the somewhat lengthy half life of ^{147}Sm (1.06×10^{11}yr), the ratio ^{143}Nd/^{144}Nd is usually measured as a fractional difference in parts per 10^4 relative to a reference chondritic reservoir ($\varepsilon_{\mathrm{Nd}}$). The systematics of Sm–Nd are very similar to those of the Rb–Sr system but because Sm and Nd are both rare earths, and hence have very similar chemical properties, both parent and daughter nuclides are considered difficult to separate one from another. Current findings suggest that they are not separated by remelting, metamorphism or sedimentary processes[8]. The early realization that the source of most magmas had a Sm–Nd ratio similar to that of chondritic meteorites suggested that the bulk earth had a chondritic composition, a feature supported by other geochemical and geophysical information[9]. As more measurements became available, small but significant differences from the bulk-earth development line were found, particularly the data from ocean-ridge basalts[10,11], which helped substantiate the steadily accumulating evidence for a chemically heterogeneous upper mantle. One extremely useful earth model attempts to integrate the isotopic results with what is known about the spatial distribution and volumetric relationships of large-scale basaltic magmatism[12]. Although Sm–Nd studies of granitoid rocks are few, the available data support the earlier interpretations of Sr and Pb isotope results from Archaean continental crust. A broad twofold division for all granitoid rocks into those of direct mantle origin and those derived from admixtures of mantle and crustal material is suggested by the isotopic data[13]. Both groups, however, seem to be temporally quite distinct. Granites younger than about 2,000 Myr contain some recycled crustal component, those older are considered primary and of direct mantle origin.

The paper by DePaolo concerns the continental evolution of part of the Proterozoic crust of the southwestern United States and bears, in greater detail, on some of these problems. DePaolo concludes from the Sm–Nd isotopic data that an early event in the Front Range of Colorado at about 1,800 Myr involved the formation of a large segment of continental crust that was derived directly from the mantle, while magmatism during the next 800 Myr, at 1,670, 1,400 and 1,000 Myr ago, involved the formation of granitic bodies from this new crustal segment. (It is interesting to note the very good agreement between the Sm–Nd and Rb–Sr isochron ages.) This story of early mantle material coupled with later crustal involvement agrees rather nicely with the picture emerging from initial ^{87}Sr/^{86}Sr ratios from granitoid plutons found in some orogenic belts of much younger age (such as the Canadian section of the Appalachian orogen[14]). A plot of initial ^{87}Sr/^{86}Sr ratios versus age shows that the older granitoid bodies have relatively low, mantle-like, initial ^{87}Sr/^{86}Sr ratios, while the younger plutons have initial ^{87}Sr/^{86}Sr ratios that are considerably higher (>0.710). The later plutons commonly occur as small, discrete bodies, are commonly peralkaline and sometimes are associated with fluorine, tin and uranium. The overall impression is that mantle-derived material was added to the crust fairly early in the orogenic cycle and that this was followed by substantially smaller amounts of material of quite different geochemical character, which became involved with, or was directly formed from, continental crust.

Of equal interest is the source of the 1,800 Myr material from the southwestern United States because it is relevant to the evolution of the mantle, in addition to placing further constraints on the use of Sm–Nd model ages. The isotopic signature of the older material, quite different from that expected of rocks derived from a chondritic-like source, suggests a depleted upper mantle with a higher Sm–Nd ratio than that of assumed undifferentiated material.

Previous ages from composites of North American continental crust tended to agree with the known geological and geochemical constraints. DePaolo argues, however, that these dates require revision: by using the Sm–Nd data from his depleted upper mantle, he brings them more into line with the known crystallization ages. Although the earlier model ages tend to be underestimated, those from Archaean composites hardly change.

Many problems raised by DePaolo's work are outside the scope of his present paper, which primarily concentrates on the Sm–Nd systematics of the Proterozoic material. One of the more intriguing, however, involves the nature of the early depletion, partly because it brings into focus the meaning of the term 'depleted mantle' and partly because it questions the role taken by this part of the mantle in later crustal generation. Geochemical variations in the mantle can be attributed to a variety of processes including melt extraction, mixing of different mantle zones or metasomatism. The implied widespread nature of DePaolo's depleted mantle suggests that it was formed by an event of planetary proportions. Whatever the nature of the early event (core ? crust ? formation) it seems to have resulted in a mantle with a higher Sm/Nd ratio than chondrites, which later either generated continental crust directly or became intimately involved with the continental

Keith Bell is in the Department of Geology, Carleton University, Ottawa.

0028-0836/81/210189-02$01.00

crust during its formation. DePaolo's recalculation of crustal-separation model ages certainly implies that large segments of continental crust were derived from a depleted mantle. The question that begs to be asked is 'What are the geochemical implications of an isotopically depleted upper mantle?'. Recent findings, for example, have shown that although initial ^{143}Nd/^{144}Nd ratios from mantle xenoliths in alkalic basalts are variable (with some isotopic ratios indicating a depleted upper mantle), certain major elements, such as CaO and Al_2O_3, are still present in sufficient amounts to generate basaltic liquids[15]. But what of typical continental crust? Can it be derived from a mantle that on the basis of the Sm–Nd data seems to be depleted? A great deal may ultimately depend on the melt volumes extracted from the mantle, spreading rates and the composition of the melts, but the question still remains as to the number of times the same part of the mantle can provide or seemingly provide 'crustal components'.

DePaolo's paper also brings out the feature, already observed from Sr isotopic data, of little or no correlation between initial isotopic composition and rock type. Data from large crustal segments, island-arc volcanics and even ocean-ridge basalts all seem to lie on or close to the depleted, Sm–Nd, mantle development line. This may suggest that most of the chemical characteristics of the various rock types were probably acquired at or close to the time of their formation. In a melting model the chemical composition of the magma would be partly controlled by the minerals present in the mantle and the partitioning of the elements between the liquid and solid phases.

Perhaps additional clues about the evolution of the continental crust can be found elsewhere. In several cases open-system behaviour in the mantle, involving volatiles and CO_2-rich fluids, has been invoked to explain some of the rare earth abundances and isotopic data[16-18]. Several papers have proposed metasomatism in the mantle and terms like 'anomalous mantle' and 'metasomatized mantle' are creeping into the literature. The idea of relating mantle metasomatism to the evolution of the continental crust may seem remote, but the mere facts that it could be considered consistent with some of the isotopic findings and that many granites, such as leucogranites and megacrystic granites, are commonly associated with deep-seated fractures suggest that mantle metasomatism in this context should not be summarily dismissed.

What sort of mantle processes affected which parent–daughter ratios, and the timing of such events, are among the thorny problems that will have to be sorted out. Even in some of the more fertile parts of the mantle, such as those from which kimberlites are derived, the U–Pb and Rb–Sr ratios are quite different from those of chondrites, suggesting that this part of the mantle may have been affected by a process that left the chondritic Sm–Nd ratio unchanged[19]. Much, of course, depends on the nature of the starting material, when it was formed, and the distribution of the chemical heterogeneities within the upper mantle. In spite of these problems there is some indication that the concentrations of primary juvenile gases and their isotopic compositions in mantle-derived materials will help in their resolution. Such information, coupled with the isotopic data from more than one system, may provide further exciting results relevant to the study of the mantle–crust system. □

---

1. Armstrong, R.L. *Rev. Geophys.* **6**, 175 (1968).
2. Hargraves, R.B. *Science* **193**, 363 (1976).
3. Hurley, P.M., Hughes, H., Faure, G., Fairbairn, H.W. & Pinson, W.H. *J. geophys. Res.* **67**, 5315 (1962).
4. Moorbath, S. *Nature* **254**, 395 (1975).
5. Gast, P.W. *J. geophys. Res.* **65**, 1287 (1960).
6. Moorbath, S. in *The Early History of the Earth* (ed. Windley, B.F.) 351 (Wiley, New York, 1976).
7. Lugmair, G.W., Scheinin, N.B. & Marti, K. *Earth planet. Sci. Lett.* **27**, 79 (1975).
8. McCulloch, M.T. & Wasserburg, G.J. *Science* **200**, 1003 (1978).
9. DePaolo, D.J. & Wasserburg, G.J. *Geophys. Res. Lett.* **3**, 249 (1976).
10. Richard, P., Shimizu, N. & Allègre, C.J. *Earth planet. Sci. Lett.* **31**, 269 (1976).
11. O'Nions, R.K., Hamilton, P.J. & Evensen, N.M. *Earth planet. Sci. Lett.* **34**, 13 (1977).
12. Wasserburg, G.J. & DePaolo, D.J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3594 (1979).
13. Allègre, C.J. & Othman, D.B. *Nature* **286**, 335 (1980).
14. Strong, D.F. *Geol. Ass. Can. Spec. Pap.* **20**, 741 (1980).
15. Jagoutz, E., Carlson, R.W. & Lugmair, G.W. *Nature* **286**, 708 (1980).
16. Menzies, M. & Rama Murthy, V. *Earth planet. Sci. Lett.* **46**, 323 (1980).
17. Wass, S.Y. & Rogers, N.W. *Geochim. cosmochim. Acta* **44**, 1811 (1980).
18. Schilling, J.-G., Bergeron, M.B. & Evans, R. *Phil Trans. R. Soc.* A **297**, 147 (1980).
19. Basu, A.R. & Tatsumoto, M. *Contr. Miner. Petrol.* **75**, 43 (1980).

Clustering of ion channels at the node of Ranvier

from L.C. Fritz and J.P. Brockes

THE segregation or clustering of surface components into discrete and functionally significant domains is an important feature of neuronal cell biology. Obvious examples are the structure of the synapse with its postsynaptic configuration containing a high density of transmitter receptors, and the segregation of various other components to dendrites, cell bodies, axons, or even local domains within these three areas. Axonal calcium channels, for instance, are largely confined to the region of the nerve terminal where they mediate transmitter release. Two systems being used to investigate the mechanisms underlying these processes of segregation are the neuromuscular junction, where acetylcholine receptors are clustered on the postsynaptic muscle fibre, and the node of Ranvier in myelinated axons of the peripheral nervous system. The node is the unensheathed area of axon membrane that lies between two successive myelin segments laid down by Schwann cells. Recent evidence suggests that the nodal membrane contains a high density of certain molecules which mediate electrical excitability and that these molecules are not detectable in the internodal membrane of the myelinated axon.

The passage of the nerve impulse depends on the presence of voltage-sensitive sodium and potassium channels, and it has long been recognized that the nodes are the sites of large inward sodium currents that mediate the saltatory conduction of the action potential[1]. More recently, Ritchie and Rogart[2] have examined the distribution of sodium channels in myelinated frog axons by comparing the equilibrium level of binding of the specific ligand tritiated saxitoxin in intact and homogenized axons. Anatomical considerations strongly suggest that the internodal axolemma should not be accessible to ligand in the intact nerve and yet the specific binding did not increase after homogenization. The authors concluded that the sodium channels are present at high density (approximately 10,000 per μm^2 assuming one toxin molecule binds to one channel) in nodes and not detectable (<25 per μm^2) in internodes. Measurements of channel gating currents have also indicated a high density at the node[3].

The localization of potassium channels is uncertain. Using voltage-clamp techniques, Chiu and Ritchie[4] have observed that significant outward current, which is blocked by tetraethylammonium ion, can be measured only after osmotic shock or collagenase digestion which presumably loosens the association of the myelin sheath with the axon. This suggests an internodal or paranodal localization of the potassium channels. In contrast, Binah and Palti[5] have claimed to detect significant potassium currents at the normal node when particular attention is paid to the holding voltage. This important issue clearly requires further investigation.

There have been several reports of morphological correlates of nodal specialization. In freeze-fracture studies, the nodal membrane has been observed to

---

L.C. Fritz and J.P. Brockes are members of the Division of Biology at Caltech.

0028-0836/81/210190-02$01.00

contain a high concentration of large intramembranous particles when compared with the internodal axolemma[6]. Furthermore, Waxman and his colleagues have shown that excitable nodes are specifically stained on the cytoplasmic surface of the membrane by a ferric ion ferrocyanide reagent (FeFCN)[7]. The suggestion that the particle distributions and staining properties of nodes indicate the presence of a high density of sodium channels has not yet been confirmed; these characteristics could relate directly to the channels or to other cytochemical features of excitable nodes.

One particularly interesting question is what role, if any, does neurone–Schwann cell interaction have in the formation of the nodal distribution of sodium channels[8]? The evidence is not decisive, but there are already some intriguing observations. Particle patches observed in freeze fracture, as well as deposition of the FeFCN stain, can be seen in developing axons before the elaboration of the myelin sheath[9,10]; the patches have also been seen in amyelinated regions of the spinal roots in the dystrophic mouse mutant[11,12]. These observations might suggest that these specializations exist in the axonal membrane before interaction with the Schwann cell. However, the Schwann cells were certainly in contact with axons in the developmental studies even if myelin had not yet been produced. On the other hand, the FeFCN stain has not been observed in amyelinated dystrophic axons[13], and it can be observed over long lengths of demyelinated axons[14]. These results might suggest that continuing axon–glial interaction is required for both the generation and maintenance of normal nodal specialization.

Nodal specialization and, in particular, the control of sodium channel clustering are not only interesting but have clinical significance in the recovery of conduction after demyelinating disease, a subject which we have not considered here. Further progress will probably be made with the development of specific cytochemical markers for sodium channels, and the exploitation of defined *in vitro* systems for resolving the nature of neurone–glial interaction. ☐

---

1. Huxley, A.F. & Stämpfli, R. *J. Physiol., Lond.* **108**, 315 (1949).
2. Ritchie, J.M. & Rogart, R.P. *Proc. natn. Acad. Sci. U.S.A.* **74**, 211 (1977).
3. Nonner, W., Rojas, E. & Stämpfli, R. *Phil. Trans. R. Soc.* **B270**, 483 (1975).
4. Chiu, S.Y. & Ritchie, J.M. *Nature* **284**, 170 (1980).
5. Binah, O. & Palti, Y. *Nature* **290**, 598 (1981).
6. Rosenbluth, J. *J. Neurocytol.* **5**, 731 (1976).
7. Waxman, S.G. & Quick, D.C. *J. Neurol. Neurosurg. Psychiat.* **40**, 379 (1977).
8. Waxman, S.G. *Trends in Neurosci.* **4**, 7 (1981).
9. Wiley-Livingston, C.A. & Ellisman, M.H. *Devl Biol.* **79**, 334 (1980).
10. Waxman, S.G. & Foster, R.E. *Proc. R. Soc.* **B209**, 441 (1980).
11. Bray, G.M., Cullen, M.J., Aguayo, A.J. & Rasminsky, M. *Neurosci. Lett.* **13**, 203 (1979).
12. Ellisman, M.H. *J. Neurocytol.* **8**, 719 (1979).
13. Waxman, S.G., Bradley, W.G. & Hartwieg, E.A. *Proc. R. Soc.* **B201**, 301 (1978).
14. Foster, R.E., Whalen, C.C. & Waxman, S.G. *Science* **210** 661 (1980).

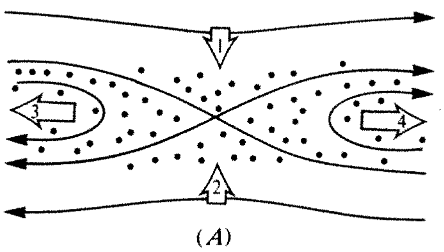
A new magnetic reconnection experiment in the laboratory

from Stanley W.H. Cowley

GREAT importance is attached to the magnetic reconnection process in plasma physics. To appreciate why this is so, it is necessary to consider the consequences of a common assumption made in theoretical plasma work, that of perfect electrical conductivity. In a medium with zero resistivity magnetic fields behave in a special way: the magnetic field lines can be considered as being 'frozen' into the medium and can be pictured as being convected along by the fluid motion. Thus, for example, the outer atmosphere of the Sun, which expands outwards at speeds ~ 400 km s^{-1}, also carries the solar magnetic field into the Solar System because this 'solar wind' gas is a highly conducting fully ionized hydrogen plasma.

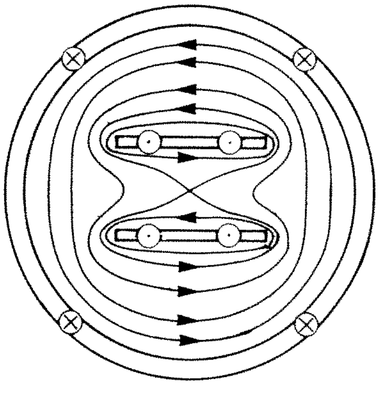
An important consequence of the perfect conductivity approximation is that plasmas and fields from separate sources can never mix. When two such plasmas and their embedded fields meet, a thin boundary layer is formed between them which carries the current required by the field jump. No field lines can link across the boundary. However, if finite resistance of the plasma is taken into account, no matter how small, 'reconnection' of the field can occur. This process is sketched in Fig. *A*, where the dotted region represents the current sheet-separated plasma regions '1' and '2' with antiparallel fields, and the solid arrows give the direction of the plasma flow. The oppositely directed fields move towards each other and 'join' at the centre of the x where the magnetic field strength is zero. The 'reconnected' lines which now link the boundary are then carried away in regions '3' and '4'. Finite plasma resistivity at the field null is required to accomplish this process, allowing the field lines to diffuse through the plasma in this region. Reconnection can therefore only occur when the perfect conductivity approximation is relaxed, although the approximation may be very good elsewhere in the flow, away from the null. Strict antiparallelism of fields '1' and '2' is by no means necessary — fields transverse to the page may be added to the figure and reconnection may still occur under very general conditions.

Reconnection is important for two main reasons. First, it changes the topology of the magnetic field and creates direct flux linkages between separated regions, along which plasmas can mix rapidly. Second, energy is liberated from the magnetic field to the plasma in the process, as indicated by the contraction of the field lines in the flow in Fig. *A*. As a consequence, the plasma in regions '3' and '4' will in general exit as a hot jet. Both aspects are of vital impor-



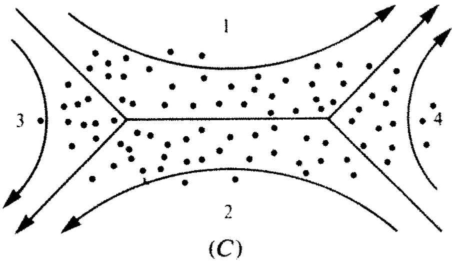
(*A*)

Sketch of the reconnection process; the lines are magnetic field lines and solid arrows indicate the direction of plasma flow. The dotted region indicates the region of large current density.



(*B*)

Cross-section through the cylindrical experiment vessel, showing the internal plates along which parallel currents are driven, returning along the vessel wall as shown. Circled dots indicate current out of the page, circled crosses current into the page.



(*C*)

Sketch of the observed field configuration at the centre of the Stenzel and Gekelman reconnection experiment during periods of imposed field increase. A neutral sheet configuration is formed rather than an x as shown in *A*.

Stanley W.H. Cowley is a SRC Advanced Fellow in the Blackett Laboratory, Imperial College, London.

0028-0836/81/210191-02$01.00

tance, for example, in the Earth's magnetosphere, where reconnection between the Earth's field and the solar wind field results in the magnetic field in the Earth's polar regions being permanently connected to the Sun. On the Sun itself explosive reconnection is believed to be responsible for solar flares, while in the laboratory an apparently similar process occurs in Tokamak 'fusion' machines and results in a disruptive instability in the plasma.

Despite considerable research, the detailed characteristics of the reconnection process have remained elusive. The problem is theoretically difficult because of the requirement to treat the region surrounding the field null, where the approximations usually used in plasma physics are invalid. This is being overcome by computer simulations, but these are, as yet, relatively crude. On the experimental front, extremely detailed plasma and field observations have been obtained from spacecraft in the Earth's magnetosphere, and although ample evidence for an 'open' magnetosphere and particle acceleration has been obtained, it has proved very difficult to pin down the details of the large-scale and temporally varying reconnection processes from single-satellite data. A final approach is to model reconnection in laboratory plasma devices. This has the advantage of repeatability and parameter control, but the detailed diagnosis of the transient plasmas whch are produced has been difficult.

Recently, the first results of a new laboratory reconnection experiment have been published (Stenzel and Gekelman *J. geophys. Res.* **86**; 649, 1981; *J. geophys. Res.* **86**; 569, 1981; Wild *et al. Phys. Rev. Lett.* **46**; 339, 1981) in which the system is sufficiently reproducible and the plasma sufficiently accessible to small internal probes that very detailed diagnostics and analysis have become possible. The experimental arrangement is shown in Fig. *B*. A steady and highly spatially uniform argon plasma is first created inside a cylindrical vessel, which is 2 m long and 1.5 m in diameter, by a pulsed hot-cathode discharge. The plasma density is ~ 10^{18} particles m^{-3}, with electron and ion temperatures of 5 and 0.5 eV respectively, and it is confined radially by a 2×10^{-3} T (20 G) axial magnetic field produced by external coils. A large current is then passed down two internal parallel aluminium plates, returning along the metal wall of the vessel, by discharging a capacitor bank. The current profile so produced resembles that of a highly damped LC circuit; it reaches a peak of $\lesssim$ 25 kA after ~ 50 μs, after which it decays and, in fact, reverses after ~ 200 μs. At any given time the magnetic field which would be produced in the absence of plasma is indicated in Fig. *B*. Typically, the field strength is ~ 10^{-2} T (~ 100 G), but is much weaker between the plates, where an x-type null occurs.

During each experiment or 'shot', field and plasma measurements are obtained at fixed positions and are recorded digitally. For each probe position up to ten shots are recorded and averaged on-line to smooth out fluctuations. The probes are then moved to new positions and the process repeated, about 300 points being so recorded in each of several transverse planes between the plates. The highly detailed data thus obtained are later analysed on a main-frame computer.

The overall behaviour of the system during the experiment is expected to depend primarily on the conductivity of the plasma, which determines the time required for the imposed magnetic field to diffuse into it. If the latter time is long compared with the characteristic time of the experiment (~ 100 μs), the plasma will behave as a perfect conductor, the field will not substantially penetrate the plasma and no reconnection will occur. Instead the plasma should be squashed down into a small volume at the centre of the chamber, leaving a vacuum around the metal conductors containing the applied field, since the energy density of the latter exceeds the initial plasma pressure by a large factor. Alternatively, if the diffusion time is short compared with the characteristic experiment time, diffusion will be dominant and the field should not differ radically from that occurring in the absence of plasma. If the expected plasma conductivity is calculated from the 'classical' formulae based on particle scattering due to particle–particle collisions, then the diffusion time turns out to be ~ 500 μs. This time is fairly long compared with the experiment time, so we expect the perfect conductivity picture to be valid, at least initially. This turns out not to be the case; instead the applied magnetic field rapidly diffuses into the plasma and reconnects across the central field null which is established. However, the field does indeed differ substantially from the vacuum field as a result of the plasma currents. Near the null, in fact, the field does not resemble an x as illustrated in Fig. *A* but instead flattens into a magnetic neutral sheet as shown in Fig. *C*. The electron current layer separating the fields is found to be ~ 10 cm wide, although narrower, ~ 1 cm structures have been observed. Such field structures remain stable for ~ 80 μs, evolving in time in response to the changing applied current. During (and due to) the decay phase of the current, however, the sheet tears up to form a magnetic island, although neutral sheet conditions are resumed after the current has been reversed in sense (at 200 μs) and starts to increase again.

Such behaviour indicates that the plasma conductivity is much lower than expected, although not so low that diffusion is totally dominant. This conclusion is confirmed in detail by the experiment diagnostics, which allow a direct measurement of the conductivity to be made, simply by dividing the measured current density by the measured electric field. It is found that the real resistivity, although highly inhomogeneous, is on average 26 times the classical value. This reduces the diffusion time to a few tens of microseconds, comparable with the time scale of variation of the applied field.

The process resulting in this 'anomolous' resistivity has yet to be identified and is the subject of continuing experiments. The enhanced electron scattering may be due to large electric field fluctuations in the gas caused by an instability associated with the plasma current. The authors point out that significant, about 20 per cent, electron density fluctuations are observed, and that the known conditions for the excitation of ion-sound turbulence are satisfied in the plasma. However, they also note that the observed resistivity does not necessarily maximize in the regions of maximum current as might have been expected, and this requires explanation.

The final result of importance concerns the electric field and the energy deposited in the plasma by the reconnection process. The total electric field is the sum of induced and potential contributions, the first arising from Faraday's law of induction and being proportional to rate of change of the magnetic field (from which it is calculated), while the second contribution is produced by slight separations of electric charge in the plasma and is measured by Langmuir probes. In this experiment, the latter very nearly cancels the former. The variation of the applied field in the reconnection experiment induces an axial electric field of ~ 100 V m^{-1}. The oppositely directed potential field, however, is found to reduce the total field to ~ 10 V m^{-1} throughout the plasma, a value which is, in fact, close to the estimate obtained with Alfvén's (1968) simple neutral sheet model. This experimental result may be of significance in magnetospheric physics, where the role of induced electric fields in particle acceleration has recently been discussed. It also resolves an apparent discrepancy of earlier experiments concerning the energy deposited in the plasma from the field. The rate at which energy is released from the field to the plasma per unit volume is simply given by the product of the current density and the total electric field. If only the induced electric field is considered, as in previous work of this nature, then the energy input estimate will be about a factor of 10 too large due to the partial cancellation effect. In this experiment the energy actually goes largely into electron heating rather than into plasma bulk flows. As a result the electron temperature rises from an initial ~ 5 eV to ~ 15 eV, the peak heating rates and temperature increases occurring at the edges of the neutral sheet rather than at its centre. Electron temperature increases are, in fact, limited to ~ 15 eV by extensive ionization of residual neutral atoms in the vessel during the early phases of the experiment ($\lesssim$ 30 μs), and by the excitation of line radiation at later times. □

ARTICLES

Neodymium isotopes in the Colorado Front Range and crust–mantle evolution in the Proterozoic

Donald J. DePaolo
Department of Earth and Space Sciences, University of California, Los Angeles, California 90024, USA

Initial $^{143}Nd/^{144}Nd$ determined for major rock types of the basement underlying the Rocky Mountains in Colorado indicate that this segment of continental crust was formed from a homogeneous and previously depleted source material in the upper mantle 1,800 Myr ago and contains no identifiable component of older crust. Subsequent magmatic events at 1,670, 1,400 and 1,000 Myr represent mainly intracrustal differentiation with small or no further additions to the crust from the mantle. These results imply that a major increase in the mass of the North American continent occurred over a narrow time interval 1,800 Myr ago. From the isotopic data an empirical model is constructed that may provide accurate crust-formation ages for rocks of all ages. Preliminary results suggest that existing estimates of crustal age distributions may require substantial revision.

UNDERSTANDING how the processes that have shaped and modified the Earth have changed through time is closely tied to the determination of the ages of rocks using radioactive isotopes. The Sm–Nd isotopic method[1–3] is particularly valuable. Because Sm and Nd have slightly different chemical properties, and because they are relatively immobile during low grade metamorphism, the radioactive decay of ^{147}Sm to ^{143}Nd, and the consequent increases in the ratio $^{143}Nd/^{144}Nd$, are reliable indications of when rock material that is now found in the Earth's crust was originally formed by the processes which create crust from the denser, more primitive materials of the mantle[4,5]. This information, which can be described as the 'crust-formation age' of a rock unit, as opposed to a crystallization age or an age of metamorphism, is crucial for determining if and how the size of the continents has changed through time.

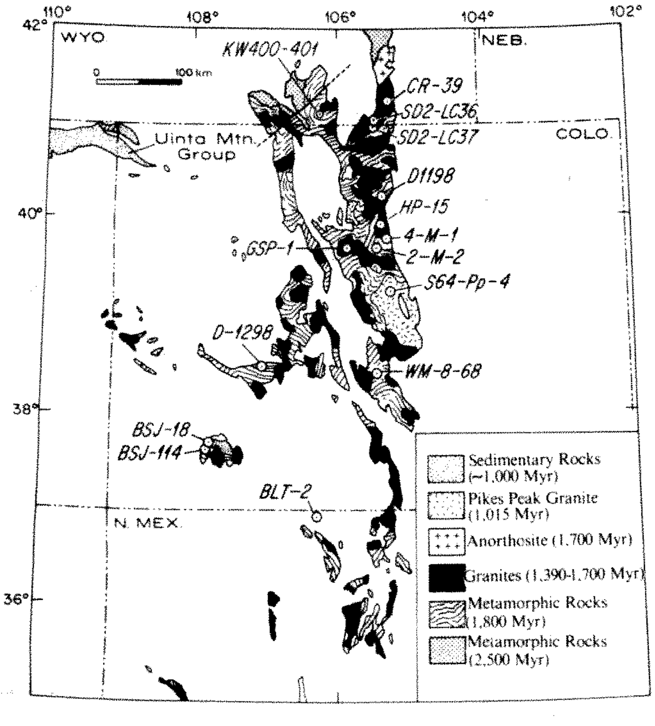


Fig. 1 Generalized basement geological map of the Southern Rocky Mountains, Colorado and New Mexico, USA, showing sample locations (adapted from ref. 11).

The definition of the Sm–Nd crust-formation age has been based on studies mainly of very old crustal rocks (≥2,500 Myr)[2,4–6]. Studies of young rocks have suggested that the model may have serious limitations for younger crust[7], but data on intermediate-aged rocks have been lacking so that the understanding of the apparent discrepancy between the oldest crust and the youngest crust has been poor. In this study, Sm–Nd isotopic measurements have been made on a previously well-characterized Precambrian basement terrain in the Front Range of Colorado[8–10], which has been dated to be 1,000–1,800 Myr old, the intermediate age range where data are lacking. This study was done to test the potential of Sm–Nd crust formation ages as well as to characterize the evolution of this continental province.

Exposures of Proterozoic Precambrian basement are extensive in Colorado in block faulted mountain ranges where Palaeozoic, Cretaceous and Tertiary uplift have exhumed the basement from beneath the Phanerozoic sediment cover (Fig. 1)[11]. The oldest rocks are gneisses which represent metamorphosed sedimentary and volcanic rocks, known as the Idaho Springs Formation[8,10]. Recent work has shown that metavolcanic rocks predominate in the northern and southern thirds of the state, whereas metapelitic rocks predominate in the central portion (ref. 12 and C. E. Hedge, personal communication). Depths of erosion have been estimated at 10–20 km; most of the rocks record an upper amphibolite facies metamorphism[13]. Three major intrusive episodes have been identified[8]. The oldest, the Boulder Creek intrusives, were associated with the regional metamorphism. They average granodiorite in composition; Rb–Sr whole rock dating has fixed the age at ~1,670 Myr. A second intrusive event at 1,400 Myr produced abundant granite with no associated regional metamorphism. The granites are of two-mica type (Silver Plume granite) in the regions where metapelites predominate in the Idaho Springs Formation, and are biotite granites (Sherman–Eolus type) in the regions dominated by metavolcanic rocks (C. E. Hedge, personal communication). The last major Precambrian intrusive event is represented by the Pikes Peak granite (1,015 Myr) which is limited to an outcrop area of ~3,400 km² at latitude 39 °N (Fig. 1). The Proterozoic terrain is bordered on the north by Archaean crust[14] and on the south by younger Proterozoic crust[15]. The contact with the Archaean terrain is a shear zone over most of its exposed length which marks a profound lithological discontinuity as well as an age province boundary[16]. The Colorado age province extends south-westwards to south-eastern California. It has been separated into two subprovinces

0028–0836/81/210193—04$01.00

of slightly different age on the basis of extensive U–Pb zircon dating[17,18]. The samples measured in this study are from the northern, slightly older subprovince.

Samples

Sample locations are shown in Fig. 1. All of the samples have been previously measured for Rb–Sr (and in some cases U–Pb) isotopes, and are currently being studied for major and trace element characteristics (ref. 8 and C. E. Hedge, personal communication). The 1,800 Myr metamorphic rocks include two samples of metavolcanic rocks from the Idaho Springs Formation in northern central Colorado and two from the time-correlative Twilight Gneiss in the San Juan Mountains ~300 km to the south-east (Table 1)[9]. Also included in the 1,800 Myr suite are two xenoliths of charnokitic granulite brought to the surface in a kimberlite diatreme near the Colorado–Wyoming border. Four samples of Boulder Creek-type intrusives have been measured, representing a range of chemical types (note Nd concentrations), and widely separated localities. The three samples of Silver Plume granite show strikingly different trace element chemistry (note Sm and Nd), which bracket the range of types found. One sample of typical Pikes Peak granite has been measured. This batholith includes a range of rock types from gabbro to granite, with granite greatly predominant[19]. One sample of the Sherman-type granite, from the type locality, has also been measured.

Analytical procedures

The procedures used for Sm and Nd isotopic measurements and the analytical uncertainties have been discussed elsewhere[20,21]. For this study the precision and accuracy of the determinations of the $^{147}Sm/^{144}Nd$ ratio are particularly important. Tracer solutions of ^{150}Nd and ^{147}Sm were calibrated using two ultrapure pieces of Sm and Nd metal (Ames Laboratory, Iowa) as gravimetric standards. Calculated concentrations for the tracers calibrated against these two standards agreed to ±0.05% for ^{147}Sm and to ±0.2% for ^{150}Nd. Relative to each other the $^{147}Sm/^{144}Nd$ ratios of the samples are precise to $\pm \leqslant 0.1\%$, but a systematic error of up to 0.25% could affect all samples. The effect on the calculated initial $^{143}Nd/^{144}Nd$ ratios measured have been normalized to be consistent with data previously reported from Caltech normalized to $^{150}Nd/^{142}Nd = 0.2096$ (Table 1). Interlaboratory comparison of Nd isotope ratios shows that $^{143}Nd/^{144}Nd$ ratios measured at UCLA should be comparable with those measured at Caltech to within ~0.002%[21], (equivalent to 0.2 ε units as defined below).

The $^{143}Nd/^{144}Nd$ ratio is also expressed here as $\varepsilon_{Nd}(T)$, the deviation from the value expected in a chondritic reservoir (CHUR), and is defined as[2]:

$$\varepsilon_{Nd}(T) = 10^4 \left[\frac{^{143}Nd/^{144}Nd_{rock}(T)}{^{143}Nd/^{144}Nd_{CHUR}(T)} - 1 \right]$$

where T is time measured backward from the present (age) and

$$^{143}Nd/^{144}Nd_{CHUR}(T) = 0.511836 - 0.1967[(\exp \lambda_{Sm} T) - 1]$$

where $\lambda_{Sm} = 6.54 \times 10^{-6}\ Myr^{-1}$ (see ref. 22). For a rock of age T, $\varepsilon_{Nd}(T)$ is the 'initial' value of ε_{Nd}; the value at the time of crystallization of the rock, and $\varepsilon_{Nd}(0)$ is the measured value.

Results

The results are given in Table 1 and Figs 2 and 3. The gneiss samples define an isochron corresponding to an age of 1,800 ± 90 Myr (2σ) (Fig. 2). These rocks are geographically remote from each other and cannot be cogenetic. There is no *a priori* reason why they should define an isochron as variations in initial $^{143}Nd/^{144}Nd$ are likely. The data therefore require that the 1,800 Myr age represents a true magmatic age within the age resolution and that the magma source which produced this crust 1,800 Myr ago was homogeneous over distances of several hundred kilometres. Note that the only types of young volcanic rocks which exhibit comparable Nd isotopic uniformity are oceanic island arc rocks[21]. Arcs located far from continental crust show a small range of ε_{Nd} worldwide (Fig. 3), whereas oceanic volcanics in general exhibit a total range from −2 to +13 (ref. 23). The data of Fig. 2 confirm the resistance of Nd isotopes to gross redistribution in crustal metamorphic processes.

The clustering of the initial ε_{Nd} values of the metamorphic rocks is shown in Fig. 3, where initial values for the plutonic rocks are also plotted. The average ε_{Nd} (1,800 Myr) of +3.7 is a

Table 1 Sm–Nd isotopic parameters for Colorado samples

Sample	Description	Sm (p.p.m.)	Nd (p.p.m.)	$^{147}Sm/^{144}Nd$*	$^{143}Nd/^{144}Nd_m$†	$\varepsilon_{Nd}(0)$†	$\varepsilon_{Nd}(T)$‡
1,800 Myr metamorphic country rocks							
4-M-1	Metarhyodacite (IS)§	7.74	33.86	0.1334	0.511281±22	−10.8	+3.8
2-M-2	Metatholeiite (IS)	3.03	9.96	0.1841	0.511903±26 (0.511884±38)	+1.3 (+0.9)	+4.2 (+3.8)
BSJ-18	Metadacite (TW)§	4.04	12.06	0.2027	0.512072±28	+4.6	+3.2
BSJ-114	Metatholeiite (TW)	3.62	11.99	0.1826	0.511869±28	+0.6	+3.9
LC-36	Charnokitic granulite	4.07	19.23	0.1281	0.511193±26	−12.6	+3.2
LC-37	Charnokitic granulite	5.29	22.62	0.1413	0.511376±18 (0.511364±30)	−9.0 (−9.2)	+3.7 (+3.5)
1,670 Myr Boulder Creek intrusives							
BLT-2	Granodiorite	3.06	15.71	0.1180	0.511157±27	−13.3	+3.5
HP-15	Granodiorite	6.06	26.55	0.1379	0.511263±32	−11.2	+1.3
WM8-68	Granodiorite	8.85	42.30	0.1266	0.511154±20	−13.3	+1.7
KW400-401	Granite	4.94	24.03	0.1243	0.511200±23 (0.511212±26)	−12.4 (−12.2)	+3.4 (+3.6)
1,430 Myr Sherman intrusives							
D1198	Granite	7.72	41.67	0.1221	0.510963±26	−17.1	−1.7
1,415 Myr Silver Plume intrusives							
D-1298	Granite (2 mica)	2.39	7.38	0.1959	0.511837±30	0.0	+0.1
CR-39	Granite (2 mica)	14.24	70.43	0.1223	0.511069±12	−15.0	+0.3
GSP-1	Granite (2 mica)	21.96	172.45	0.0770	0.510605±22	−24.1	−2.5
1,015 Myr Pikes Peak batholith							
S64-Pp4	Granite	18.36	104.62	0.1061	0.511170±24	−13.0	−1.2

* Analytical precision ±0.1%; systematic error due to tracer calibration uncertainties is limited to ±0.25%.

† Measured values, with uncertainty at 95% confidence, normalized to $^{146}Nd/^{142}Nd = 0.63613$. Replicate measurement given in parentheses.

‡ Initial values; uncertainties typically ±0.5 ε units.

§ IS, Idaho Springs formation; TW, Twilight Gneiss.

distinctly mantle-type value. Shown schematically is the typical value of ε_{Nd} (1,800 Myr) expected in the 2,700 Myr crust which borders the Colorado province to the north[24]. The ε_{Nd} values of 1,800 Myr gneisses give no evidence of influence from older crust. The data require that the Colorado crust be derived from the mantle *in toto* 1,800 Myr ago with the crust formation process involving no reconstitution of older crust.

The line $\varepsilon_{Nd}=0$ (Fig. 3) has been interpreted as the Nd isotopic evolution of the bulk Earth, or equivalently, of primitive mantle[4]. The initial ε_{Nd} of the 1,800 Myr samples demonstrates that this crust was not derived from primitive mantle, but from mantle which had previously acquired an elevated Sm/Nd ratio, a characteristic of parts of the mantle that have been depleted in 'crustal components'. The Colorado crust, therefore, seems to have been derived from depleted mantle, presumably part of the mantle that had been involved in producing crust during the Archaean. Based on data from island arcs[25] these conclusions also apply in general to crust being formed from the mantle at present, suggesting that the 'mantle magma source' vector shown in Fig. 3 may be a good approximation to the Nd isotopic evolution of the mantle source of continental crust over the past 2,000 Myr.

In contrast to the gneisses, the Boulder Creek intrusives show a small but distinct range of initial ε_{Nd} values ranging from the inferred mantle value near +3.7 down to +1.3. The Silver Plume and Sherman granite samples have ε_{Nd} values that are still lower, +0.1 to −2.5, and are far displaced from the mantle ε_{Nd} curve (Fig. 3). The Pikes Peak sample also has low ε_{Nd}, although it is not lower than those of the Silver Plume samples.

When all of the data on the intrusives are considered, there is a strong correlation between decreasing initial ε_{Nd} values and increasing time elapsed since the crust-forming event at 1,800 Myr. Furthermore, the ε_{Nd} values fall within or slightly below the approximate average ε_{Nd} evolution zone of the 1,800 Myr gneisses (Fig. 3). These observations strongly suggest that the intrusive events at 1,670, 1,400 and 1,000 Myr involved mainly melting of the 1,800 Myr crust rather than derivation of magmas from the mantle. However, two of the Boulder Creek samples have initial ε_{Nd} values identical to the inferred mantle value. This could indicate the presence of mantle-derived magma at 1,670 Myr, but would also be consistent with some of the Boulder Creek magmas being formed by melting of the high-Sm/Nd tholeiitic-type 1,800 Myr metabasaltic rocks such as samples 2-M-2 and BSJ-114. For the Silver Plume intrusives, the conclusion based on the isotopic data that they are crustally-derived, is perfectly compatible with their peraluminous bulk chemistry, which independently argues for derivation from the crust. The Sherman intrusives have less aluminous compositions, yet the initial ε_{Nd} of the sample measured is identical to the Silver Plume samples. This datum suggests that the Sherman-type granite was also crustally derived; the different bulk chemistry being presumably a reflection of the igneous, rather than sedimentary, character of the crust in northern Colorado.

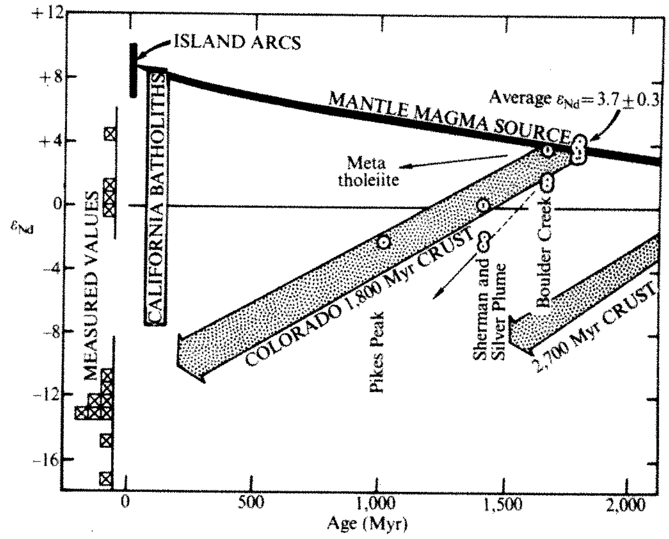


Fig. 3 Measured ε_{Nd} values are shown in histogram format at left (GSP-1 not shown). Large arrows show ε_{Nd} evolution for typical Colorado crust, and typical 2,700 Myr crust. The curve labelled 'mantle magma source' connects the average ε_{Nd} of the 1,800 Myr samples with the average for modern oceanic island arcs. Small arrows give the ε_{Nd} evolution for typical 1,800 Myr metatholeiitic basalt and the Silver Plume sample (GSP-1) with lowest initial ε_{Nd}. The Colorado crust seems to have been derived solely from the mantle at 1,800 Myr in contrast to the Mesozoic batholiths of California, which are mixtures of mantle-derived components and older crust[7].

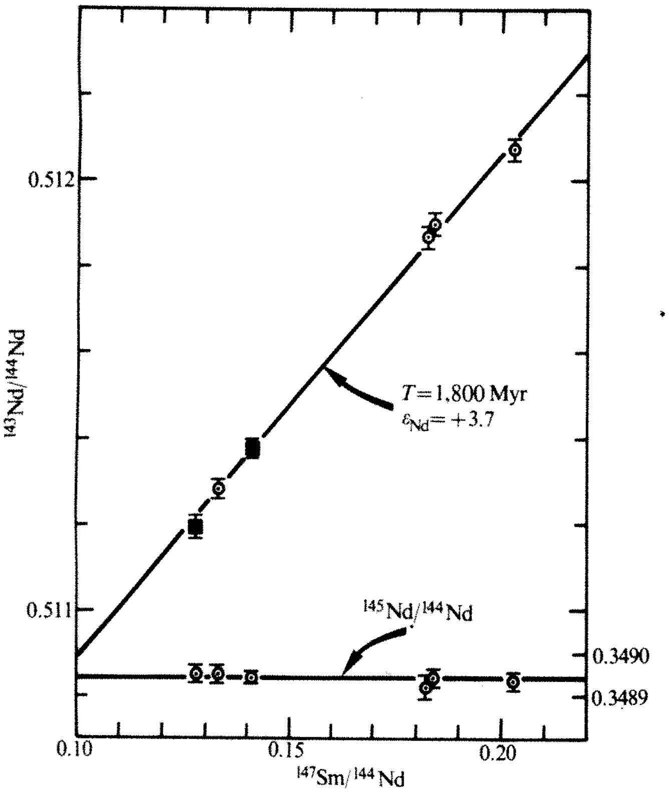


Fig. 2 Sm–Nd isotopic evolution diagram showing data for metamorphic rock samples from the Idaho Springs Formation and the Twilight Gneiss. All six data points lie within experimental uncertainty (95% confidence) of a line whose slope corresponds to an age of 1,800±90 Myr. Also shown are the measured ^{145}Nd/^{144}Nd ratios for the same samples, with the scale shown at the right. This ratio is invariant in natural samples and is measured in this laboratory as an internal check on the measurement techniques for ^{143}Nd/^{144}Nd. The measured values of ^{145}Nd/^{144}Nd for the six samples lie within error limits of the average value of 0.348945 measured for all samples to date. ⊙, Metavolcanic; ■, granulite xenoliths.

The interpretation of the Pikes Peak datum is more equivocal, as the alkalic composition of that intrusion represents a fundamentally different magma type[19]. It is possible that the Pikes Peak parent magma was derived from a mantle reservoir that had ε_{Nd} different from that shown in Fig. 3. However, petrological and other isotopic studies[19,25] indicate that the Pikes Peak granite represents the end product of a complex magmatic history of crystallization and crustal assimilation, beginning with a mantle-derived basaltic magma. This would leave the granite with a distinctly crustal isotopic composition even though the magmatism originated in the mantle. Further data on the more mafic lithologies of the Pikes Peak intrusion will be necessary to determine the Nd isotopic nature of the source of the parent basalt. The problem of mantle-derived magmas which assimilate copious amounts of continental crust has not been dealt with sufficiently and could also be a factor in the ultimate interpretation of the Silver Plume and Boulder Creek samples.

Conclusions

The present data suggest that the continental crust of Colorado was formed by differentiation from the mantle over a relatively narrow time interval at 1,800 ± 100 Myr ago. Subsequently little material was added to the crust from the mantle, with the possible exception of the Pikes Peak event at 1,000 Myr. The production of the Colorado continental crust was apparently not a continuing process, but a fairly sharply defined event. These conclusions, which could apply to the basement of much of the southwestern United States, in conjunction with previous findings[5], suggest a strong episodicity in the evolution of North America with regard to crust-forming events; crystallization and metamorphic ages show that tectonism and magmatism are more likely to be continuous, or at least to occur more frequently than major crust-forming events. The episodicity in crust production is difficult to understand in the context of current models of plate tectonics, which imply continuous crust production. More data are needed especially from other continents as an individual continent may not preserve a representative history of the crust-forming process.

The possibly well-defined evolution of $\varepsilon_{\mathrm{Nd}}$ in a mantle reservoir which has acted as the source of juvenile crust through the past 2,000 Myr has implications for the dating of crust-forming events and for crust–mantle evolution models. It had been suggested[4] that crust was derived from mantle reservoirs with $\varepsilon_{\mathrm{Nd}} = 0$ and this model was used to calculate[5] crust-formation ages, T_{CHUR}. However, as shown here, crust is not formed with $\varepsilon_{\mathrm{Nd}} = 0$ but from depleted mantle with $\varepsilon_{\mathrm{Nd}} > 0$. Hence T_{CHUR} does not give a reliable crust-formation age. Instead, crust-formation ages should be calculated relative to the mantle evolution curve shown in Fig. 3, which can be described by $\varepsilon_{\mathrm{Nd}}(T) = 0.25\ T^2 - 3T + 8.5$, where T is age. As mantle having this $\varepsilon_{\mathrm{Nd}}$ evolution can be assumed to be depleted with respect to so-called incompatible elements, we denote this revised model age T_{DM} (depleted mantle). As an example of the difference, the T_{CHUR} age of the Colorado samples range from 1,200 to 1,550 Myr, averaging about 1,350 Myr, for the samples with relatively low Sm/Nd. T_{DM} ages are 1,600–1,900 Myr, averaging 1,800 Myr. T_{CHUR} ages underestimate the true crustal age. T_{DM} ages for the Colorado rocks are relatively accurate despite crystallization ages that vary by 800 Myr. This confirms that the fractionated Sm/Nd ratios of crustal rocks relative to the mantle date mostly from the time of original crust formation.

When the crust formation ages of ref. 5 are revised to T_{DM}, calculated ages are plausible and consistent with known crystallization ages. For example, the T_{DM} age of a Grenville composite is 1,300 Myr, identical to the T_{DM} age of a granite from the Llano uplift of Texas, which has been interpreted to be part of the same province. T_{DM} of the North American shale composite becomes 1,750 Myr. Archaean ages change little due to offsetting errors related to revision of the CHUR curve[22]. T_{DM} ages provide a self-consistent, empirical model for evaluating the formation times of continental crust. This study and previous published Sm–Nd results suggest that earlier estimates of crustal age distributions based on Rb–Sr and K–Ar data[26] may have erred by underestimating the amount of old crust. More precise estimates should be possible using T_{DM} when more data become available.

The initial value of $\varepsilon_{\mathrm{Nd}}$ for 1,800 Myr crust could be accommodated for a model of a growing continental crust with no crustal recycling or for a model of a constant-mass crust with decreasing crustal recycling rates towards the present[27,28]. In either case it is required that the amount of mantle which interacts with the crust be limited to a small fraction (≤ 0.3) of the total mantle. This result supports a two-layered mantle model[28–30] and suggests that it applies over most of Earth history. The implication is that the basic tectonic–geochemical cycles involving the crust and mantle have not changed from the late Archaean to the present except possibly with respect to rates.

I thank Carl E. Hedge for well-documented samples and Tom Owens for assistance with analysis. This investigation was supported by NSF grant EAR78-12966.

Received 4 December 1980; accepted 11 March 1981.

1. Lugmair, G. W. *Meteoritics* **9,** 369 (1974).
2. DePaolo, D. J. & Wasserburg, G. J. *Geophys. Res. Lett.* **3,** 249 (1976).
3. Richard, P., Shimizu, N. & Allegre, C. J. *Earth planet Sci. Lett.* **31,** 269 (1976).
4. DePaolo, D. J. & Wasserburg, G. J. *Geophys. Res. Lett.* **3,** 743 (1976).
5. McCulloch, M. T. & Wasserburg, G. J. *Science* **200,** 1003 (1978).
6. Hamilton, P. J., Evensen, N. M., O'Nions, R. K., & Tarney, J. *Nature* **277,** 25 (1979).
7. DePaolo, D. J. *Science* **209,** 684 (1980).
8. Peterman, Z. E. & Hedge, C. E. *Can. J. Earth Sci.* **5,** 749 (1967).
9. Barker, F., Arth, J. G., Peterman, Z. E. & Friedman, I. *Bull. geol. Soc. Am.* **87,** 189 (1976).
10. Epis, R. C. & Weimer, R. J. (eds) *Studies in Colorado Field Geology* Vol. 8 (Colorado School of Mines, 1976).
11. King, P. B. *The Evolution of North America* (Princeton University Press, 1977).
12. Tweto, O. *Misc. Field Studies Map MF*-788 (US Geol. Survey, 1976).
13. Hutchinson, R. M. in *Studies in Colorado Field Geology* Vol. 8, 32 (Colorado School of Mines, 1976).
14. Peterman, Z. E. *Econ. Geol.* **74,** 1544 (1979).
15. Wasserburg, G. J. in *Advances in Earth Science* (MIT Press, 1966).
16. Condie, D. C. in *The Early History of the Earth* (ed. Windley, B. F.) 499 (Wiley, New York, 1976).
17. Silver, L. T. & Anderson, T. H. *Geol. Soc. Am. Abstr. Prog.* **6,** 955 (1974).
18. Silver, L. T. in *Continental Tectonics*, 26 (National Academy of Sciences, Washington DC, 1980).
19. Barker, F., Wones, D. R., Sharp, W. N. & Desborough, G. A. *Precambr. Res.* **2,** 97 (1975).
20. Papanastassiou, D. A., DePaolo, D. J. & Wasserburg, G. J. *Proc. 8th Lunar Sci. Conf.* 1639 (1977).
21. DePaolo, D. J. *J. geophys. Res.* (in the press).
22. Jacobsen, S. B. & Wasserburg, G. J. *Lunar and Planetary* **11,** 502 (1979).
23. O'Nions, R. K., Hamilton, P. J. & Evensen, N. M. *Earth planet Sci. Lett.* **34,** 13 (1977).
24. DePaolo, D. J. & Wasserburg, G. J. *Geochim. cosmochim Acta* **43,** 615 (1979).
25. DePaolo, D. J. & Johnson, R. W. *Contr. Miner. Petrol.* **70,** 367 (1979).
26. Hurley, P. M. & Rand, J. R. *Science* **164,** 1229 (1969).
27. DePaolo, D. J. *Geochim. cosmochim. Acta* **44,** 1185 (1980).
28. O'Nions, R. K., Evensen, N. M. & Hamilton, P. J. *J. geophys. Res.* **84,** 6091 (1979).
29. Wasserburg, G. J. & DePaolo, D. J. *Proc. natn. Acad. Sci.* **76,** 3594 (1979).
30. Sun, S. S. & Hanson, G. N. *Geology* **3,** 297 (1975).

Biological export of shelf carbon is a sink of the global CO_2 cycle

John J. Walsh[*], Gilbert T. Rowe[*], Richard L. Iverson[†] & C. Peter McRoy[‡]

[*] Oceanographic Sciences Division, Brookhaven National Laboratory, Upton, New York 11973, USA
[†] Department of Oceanography, Florida State University, Tallahassee, Florida 32306, USA
[‡] Institute of Marine Sciences, University of Alaska, College, Alaska 99701, USA

Measurements of carbon metabolism, production and exchange along food webs suggest that large fractions of the organic matter produced on continental shelves must be exported to continental slopes. The annual loss of organic matter from continental shelf ecosystems is far greater than in the open ocean. If part of the loss of nearshore primary production has increased in those coastal zones where anthropogenic inorganic nutrient supplies have been consistently increasing since the industrial revolution, then burial and diagenesis of this material in slope depocentres could represent the 'missing BMTs of carbon' in global CO_2 *budgets.*

0028–0836/81/210196—06$01.00

THE fate of anthropogenic CO_2 emissions associated with the industrial revolution is not fully known[1]. We cannot specify either the steady-state annual fluxes of carbon in a global CO_2 budget or the changes among the major reservoirs of CO_2 (atmosphere, land and ocean) over the past century. At a rate of increase in world fuel consumption of 4.3% yr^{-1}, a doubling of the CO_2 content of the atmosphere could lead to a 'greenhouse' effect[2], with a concomitant 2–3 °C increase in the ocean's temperature by 2035, similar to the temperature increase since the last glaciation.

The role of the marine biota as a carbon sink has been thought insignificant in the global CO_2 cycle because of nutrient limitation of carbon fixation in the open ocean[3]. We propose here, however, that a large amount of the CO_2 fixed each year during nutrient-unlimited photosynthesis on the continental shelves is apparently deposited as phytodetrital particles with a C/N weight ratio of ~5/1 on the adjacent continental slope. After return of nitrogen to the water column during early diagenesis, the refractory carbon particles in the sediment, with a C/N weight ratio of ~10/1, represent a major organic sink of the global carbon cycle. Part of the nitrogen source for the original coastal phytodetritus might even be manipulated to influence future CO_2 increases of the atmosphere.

Unbalanced global carbon budgets

Of the 5.2×10^9 tons C yr^{-1} now emitted from consumption of fossil fuel and cement production, about half[4–6] remains in the carbon pool of the atmosphere (Table 1). The equilibrium between the partial pressure of CO_2 in the atmosphere and that of the upper ocean is such that atmospheric CO_2 usually penetrates the upper mixed layer (~75 m) of the sea. During the spring bloom in high latitudes, for example, a primary production rate of only 1–2 g C m^{-2} day^{-1} so far exceeds the atmospheric CO_2 invasion rate that the upper 60 m becomes seasonally depleted of total CO_2 (D. Hood and L. Codispoti, personal communication). The photosynthetic and biogenic $CaCO_3$ sinks in world-wide surface waters causes the average total inorganic carbon concentration to increase with depth to a maximum at ~1,500 m, where a CO_2 source occurs within the O_2 minimum layer[7].

Because of this vertical gradient, it is not clear that a significant fraction of the CO_2 of the upper ocean actually diffuses down across the main thermocline, but surface CO_2 reaches the deeper layers due to bottom water formation at the polar boundaries. The atmospheric CO_2 which penetrates to the deep sea[8], and that formed by dissolution of the carbonate sediments beneath the main lysocline at ~3.7 km is stored as HCO_3^-. In this way ~37% (ref. 3) of the CO_2 emitted from burning fossil fuel, or ~2.0×10^9 tons C yr^{-1}, may be sequestered as HCO_3^-. Such a polar accumulation of CO_2, however, is offset to some unknown degree by outgassing of CO_2 at the air–sea interface in the equatorial regions. Without consideration to any other source terms of CO_2 (Table 1) ~0.65×10^9 tons C yr^{-1} is still 'missing' from such a chemical global budget.

The removal of land biota in undeveloped regions since 1750 may have been a source of CO_2 to the atmosphere. Changing patterns of deforestation mean, however, that the boreal terrestrial ecosystem[9] may now be a sink of CO_2, while the tropical forests[10] may be a recent source. How much terrestrial carbon fluxes to the atmosphere or ocean is debatable[11,12], but recent estimates suggest a present land source of $1\text{–}2\times10^9$ tons C yr^{-1} (Table 2). Part of this carbon loss from land could be stored in the intermediate form of both dissolved organic carbon (DOC) and particulate detrital carbon leached into riverine systems. River systems may release as much as 0.8×10^9 tons DOC yr^{-1} to the shelf while sediment accumulation of terrigenous particulate matter[14] within estuarine systems has been estimated[15] to be 0.2×10^9 tons C yr^{-1}. If the land is a source of CO_2, at least 1.55×10^9 tons C yr^{-1} could be 'missing' in a steady-state global carbon budget (Table 2) that includes fossil fuel and land source terms; atmospheric, DOC, estuarine sediment and deep ocean chemical sinks; but ignores the marine biota as a possible sink.

Table 1 An unbalanced chemical budget of the global CO_2 cycle ($\times10^9$ tons C yr^{-1})

Sources	Fossil fuel emission	5.20
	Equatorial outgassing	?
	Volcanism	0.05
Sinks	Carbonate neutralization	2.00
	Atmospheric storage	2.60
Missing carbon		0.65

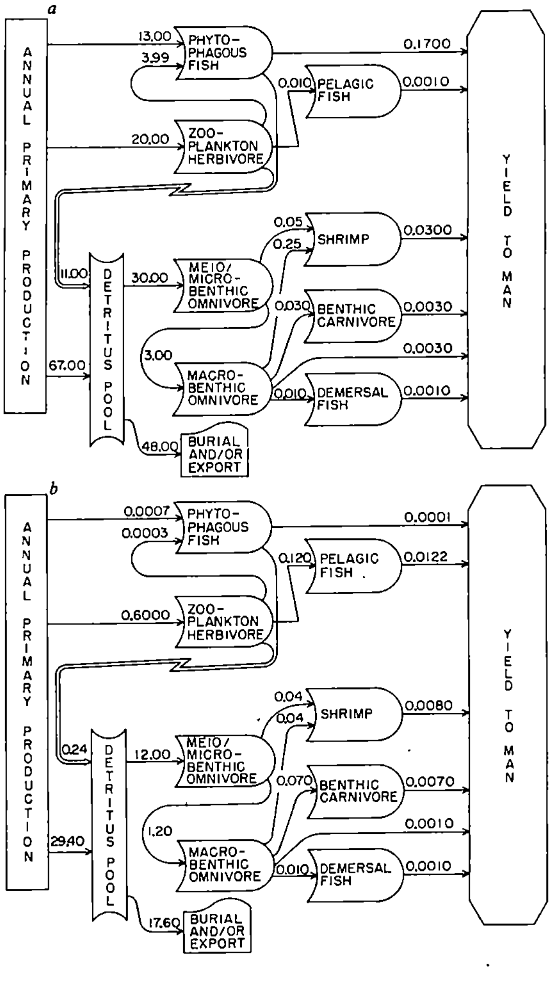


Fig. 1 A carbon budget (g C m^{-2} yr^{-1}) of yield to man derived mainly from riverine input of nitrogen on the: *a*, Louisiana–Texas; and *b*, West Florida shelves of the Gulf of Mexico.

Marine detrital sink

A marine primary production[11,15] of ~26×10^9 tons C yr^{-1} can be partitioned into 20×10^9 tons C yr^{-1} derived from the open ocean (~3×10^8 km^2) and 6×10^9 tons C yr^{-1} from the continental shelves (~3×10^7 km^2). Over 90% of the photosynthetic uptake of CO_2 in the open ocean is grazed within the upper 100 m, however, with most of the phytoplankton carbon and nitrogen recycled by herbivores as CO_2 from respiration and as ammonium from excretion. ^{15}N estimates of the daily nitrogen demand of oceanic primary production suggest[16] that at least 90% is met by the recycled NH_4^+. Those organic detrital particles, which are not grazed and sink out of the euphotic zone (~100 m) of the open ocean, amount to ~2×10^9 tons C yr^{-1} metabolized at depths down to 1,000–2,000 m, where the CO_2 produced is sequestered within the bicarbonate cycle. Heterotrophic metabolism within the O_2 minimum layer poises the diffusion gradient of total CO_2 towards the surface, that is the diffusive flux of CO_2 should be towards the euphotic zone[7]. Because CO_2 returns to the euphotic zone with the 'new' nitrate diffusing upwards to support primary production, the importance of open ocean organic detritus as a net sink for CO_2 is probably small[16]. Of the minute amount (~1% of the open production) of particles that reach 4,000–5,000 m, a fraction of this organic matter with a C/N ratio ≥10 is finally incorporated (Table 2) in bottom sediments of a similar C/N ratio[17].

Phytoplankton populations of the continental shelf, in contrast to those of the open ocean, do not seem to be either nigrogen-limited or grazer-controlled for parts of the year but rather light-limited as evidenced by spatial and temporal outbreaks of algal blooms with C/N ratios ≃5 in coastal waters[14,18,19]. In further contrast to the oligotrophic open ocean, ^{15}N estimates of nitrogen uptake by coastal phytoplankton suggest that at least 50% of the daily nitrogen demand of photosynthesis is met in the form of nitrate off Peru, western Mexico[20], California[21], in the Middle Atlantic Bight[22], in the antarctic[23], and in the south-east Bering Sea[24]. The major sources of nitrate in these systems seem to be river runoff[25] and cross-shelf exchange of subsurface slope water onto the shelf[24,26,27]; the importance of nitrification on the shelf is unknown.

After the decline of the Peruvian anchoveta fishery, removal of fish by man from each shelf amounts to <1% of the annual primary production[28]. Because nitrate supply is considered an estimate of available 'new' production[16,29], because fish yield is not a significant particulate nitrogen loss, and because plankton biomass is not continuously increasing on the continental shelf, ~50% of the 'new' primary production should at times be exported from the shelf water column each year as opposed to the <10% loss derived from 'new' production of the open ocean euphotic zone[16]. Annual carbon budgets for the shelf food web supported by riverine input of nitrogen in the Gulf of Mexico (Fig. 1), from shelf-break input within the Bering Sea (Fig. 2) and the Peru upwelling ecosystem[28], and from the more complex sources of nitrogen in the New York Bight[30], all indicate that 50% of the primary production is apparently not consumed by shelf organisms at 20, 30 or 60° latitudes.

Table 2 An unbalanced biological budget of the global CO_2 cycle ($\times 10^9$ tons C yr^{-1})

Sources	Fossil fuel emission	5.20
	Equatorial outgassing	?
	Volcanism	0.05
	Deforestation	2.00
Sinks	Atmospheric storage	2.60
	Riverine release of DOC	0.80
	Terrestrial sediment	0.20
	Open ocean detritus	0.10
	Carbonate neutralization	2.00
Missing carbon		1.55

A balanced budget of the carbon requirements of the demersal food web on the Scotian shelf[31] was obtained using net production/biomass ratios and transfer efficiencies similar to those above. However, this estimated production, 102 g C m^{-2} yr^{-1}, did not include the spring bloom which amounts to 40–50% of the yearly production within the adjacent Mid-Atlantic Bight. The similarity between the cross-shelf nitrate flux and fish yield of the Nova Scotian and Bering shelves[24] suggests that the annual primary production may be the same. Using an estimate of primary production of at least 200 g C m^{-2} yr^{-1}, we find a budget for the Nova Scotian shelf could not account for 50% of the annual production. Furthermore, a phytoplankton cell, falling to the bottom of the Scotian shelf, might be entrained in the subsurface offshore flow near the shelf-break[32]; sediment chlorophyll content to 800 m on the upper slope is, in fact, 30 times greater than on either the Scotian shelf or the lower slope[31].

Offshore flows of water also occur during upwelling off Peru and on estuarine-type shelves such as the New York Bight and the Bering Sea, suggesting that cross-shelf transport of biogenic particles is a ubiquitous phenomenon, perhaps intensified as a result of winter storms. Deposition of organic carbon in shelf sediments has been suggested[15,31] as a sink in the global carbon

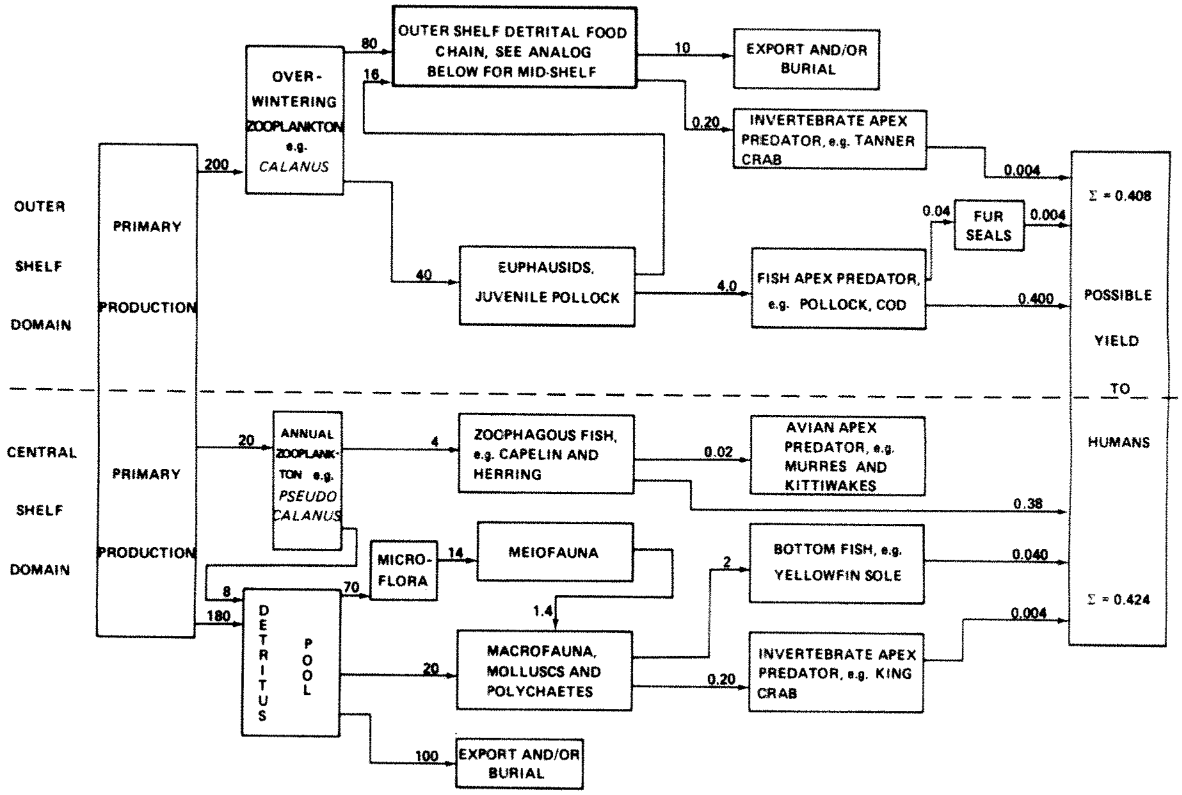

OUTER SHELF DOMAIN
PRIMARY PRODUCTION
200
OVER-WINTERING ZOOPLANKTON e.g. CALANUS
80
16
OUTER SHELF DETRITAL FOOD CHAIN, SEE ANALOG BELOW FOR MID-SHELF
10
EXPORT AND/OR BURIAL
0.20
INVERTEBRATE APEX PREDATOR, e.g. TANNER CRAB
0.004
40
EUPHAUSIDS, JUVENILE POLLOCK
4.0
FISH APEX PREDATOR, e.g. POLLOCK, COD
0.04
FUR SEALS
0.004
0.400
Σ = 0.408
POSSIBLE YIELD TO HUMANS
CENTRAL SHELF DOMAIN
PRIMARY PRODUCTION
20
ANNUAL ZOOPLANKTON e.g. PSEUDO CALANUS
4
ZOOPHAGOUS FISH, e.g. CAPELIN AND HERRING
0.02
AVIAN APEX PREDATOR, e.g. MURRES AND KITTIWAKES
0.38
MICRO-FLORA
14
MEIOFAUNA
8
70
180
DETRITUS POOL
1.4
2
BOTTOM FISH, e.g. YELLOWFIN SOLE
0.040
0.004
Σ = 0.424
20
MACROFAUNA, MOLLUSCS AND POLYCHAETES
0.20
INVERTEBRATE APEX PREDATOR, e.g. KING CRAB
100
EXPORT AND/OR BURIAL

Fig. 2 A carbon budget (g C m^{-2} yr^{-1}) of yield to man derived mainly from the shelf-break input of nitrogen within the outer and central shelf domains of the Bering Sea.

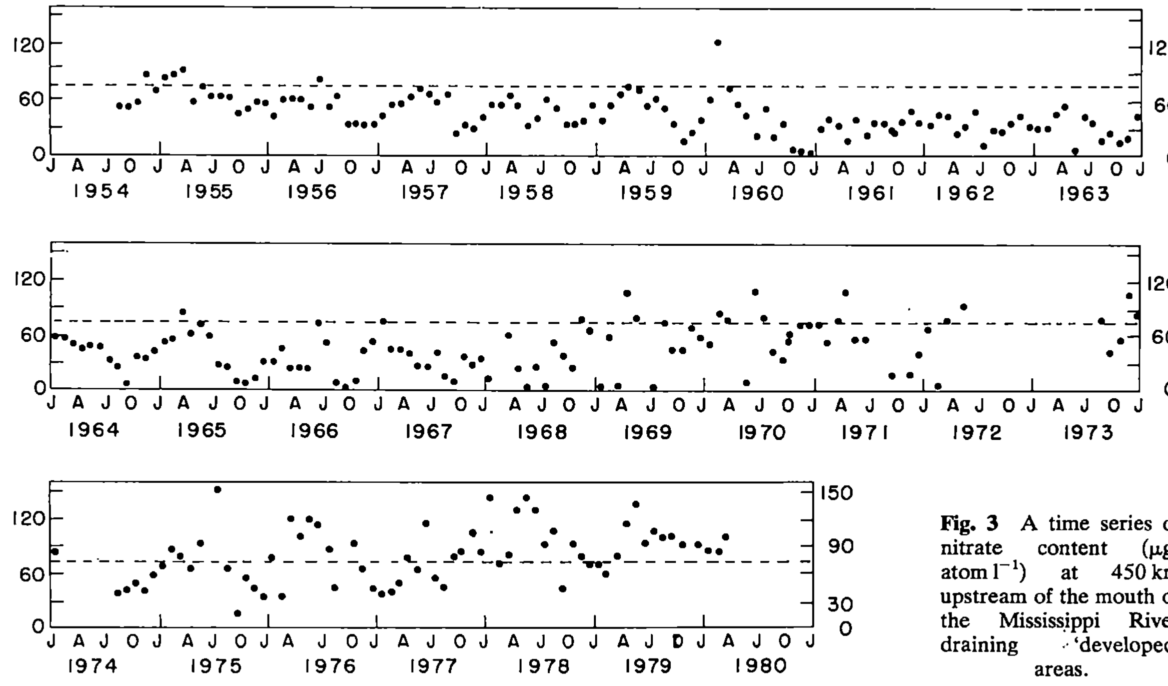


Fig. 3 A time series of nitrate content (μg-atom l^{-1}) at 450 km upstream of the mouth of the Mississippi River draining 'developed' areas.

cycle, but most recent shelves have sediments consisting of relict sands with <0.5% carbon content. However, surficial sediments of the upper slopes, for example, off the US east coast and the Gulf of Mexico, contain low C/N ratios (~5) with organic carbon concentrations >1%, implying marine origin[14]. After early diagenesis, the C/N ratio of particulate matter in these sediments increases to ≥10. That is for every 5 g of carbon and 1 g of nitrogen returned to the slope water column, a maximum of 5 g C could be left behind in the sediments. A carbon export of 3×10^9 tons C yr^{-1} from the total shelf production could now be deposited[1], with perhaps 1.5×10^9 tons C yr^{-1} accumulated[14], within slope depocentres, constituting the sink of missing carbon (Table 2) in the global CO_2 budget. Estimates of the annual detrital nitrogen export from some of these shelves are balanced by independent budgets of shelf nitrogen recycling and of the return flux of nitrate from offshore and riverine sources[24,28,30].

Transients in anthropogenic nitrogen loading

Most of the CO_2 produced by burning fossil fuel has been introduced into the atmosphere within the past two decades[5,6]; before the 1950s most of the emitted CO_2 was thought to have been stored in the ocean[34]. Presumably a transient in one or more of the sinks of CO_2 (Table 2) must have occurred in response to this increase of the fossil fuel source. However, a decline in temperate forest emission[9] of CO_2 from a source of $\sim 3 \times 10^9$ tons C yr^{-1} in 1850 to a net sink after 1950 may have been partially balanced by the increase in fossil fuel emission. For example, ignoring both marine photosynthesis as a sink and the land as a source term, a budget[35] of chemical storage of CO_2 within the deep ocean from 1850 to 1950 may have overestimated[36] the initial carbon content of the pre-industrial atmosphere by $\sim 50 \times 10^9$ tons C. In this case, at least 0.5×10^9 tons C yr^{-1} could have been lost to an unknown biotic sink without considering the changes in deforestation during most of this century.

Table 3 Major river discharge ($\times 10^3$ m^3 s^{-1}) and nitrate content (mg-atom NO_3 m^{-3}) in relation to human population centres (1970 census)

Oligotrophic N sources (5.5×10^8 people) Africa–South America–Oceania				Eutrophic N sources (30.1×10^8 people) North America–Europe–Asia			
River	Flow	NO_3	Ref.	River	Flow	NO_3	Ref.
Amazon	180	10	53	Yantze	34	70	*
Congo	41	7	54	Mississippi	18	150	†
Orinoco	20	8	*	Mekong	11	75	55
Niger	6	8	54	Rhine	2	350	43

Personal communications: * J. Edmond; † C. Demas.

Other analyses of abiotic storage of CO_2 in the ocean by vertical mixing[37] and/or polar sinking[38] over the past few decades neglect biological fixation of carbon in the sea and cannot account for all the CO_2 recently emitted by fossil fuel burning. The open ocean photosynthetic loss from the water column is unlikely to have increased[16] as a result of either natural or anthropogenic nutrient loading[18]. The shelf ecosystems respond to seasonal nutrient enrichment, however, when light limitation ceases, by formation of spring phytoplankton blooms[18], which seem to be exported to the slope depocentres. In summer stratified conditions of nitrogen depletion, anthropogenic nutrient addition to surface coastal waters would also increase the yield or export from this ecosystem. Poor records of primary production in the ocean on the decadal time scale[39] mean that it is difficult to demonstrate the eutrophication of coastal waters by direct observation. The daily production of the mid-shelf spring bloom in the New York Bight has apparently not changed, for example, over the past 25 yr [27], but spatial sampling was inadequate to determine whether nearshore production had increased. The Baltic[40] and Waddensea are exceptions, where the data base indicates that annual production of the nearshore zone has apparently increased from 80 to 240 g C m^{-2} yr^{-1} over the past 20 yr [41].

The lack of change in nutrient content at midshelf off New York, away from the influence of the Hudson River over the past 25 yr [27], suggests that only coastal anthropogenic sources of nutrients, not shelf-break sources, could have reasonably changed in response to the industrial revolution. The phosphate content[42] of the Rhine River has, in fact, increased 10-fold over the past 50 yr. The nitrogen content of the Rhine[43] almost doubled between 1959 and 1968, while that of the Seine (Y. Monbet, personal communication) increased fourfold to ~400 μg-atom N l^{-1} from 1965 to 1975. The mean winter nitrate content of the North Sea shelf is also at least two- to fourfold higher off the Thames estuary, the Wash and the Rhine estuary than at the edge of the shelf[44]. Indeed, as much as 350 μg-atom NO_3 l^{-1}, ~10-fold the concentration of NO_3 in deep slope water, has been found both in the Scheldt estuary and 10 km off the Belgian coast[45].

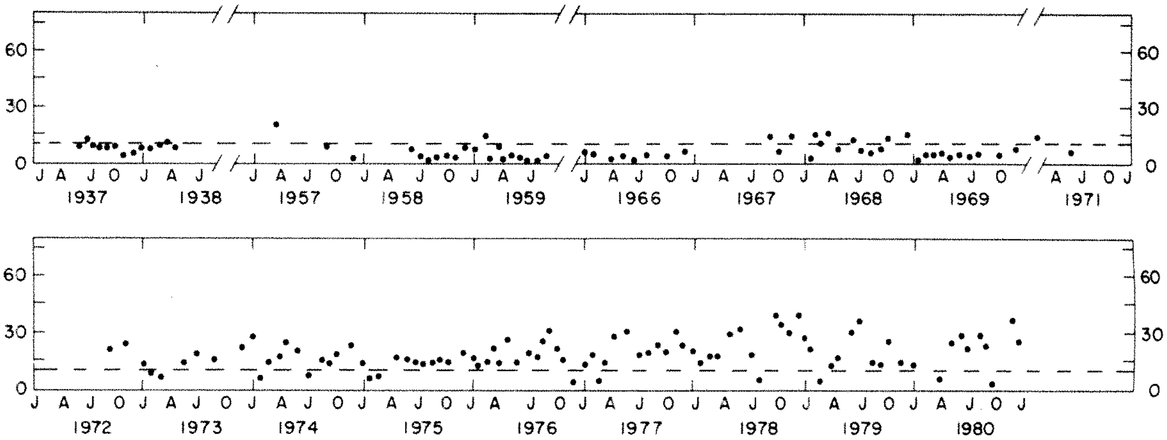


Fig. 4 A time series of the nitrate content (µg-atom l^{-1}) at 50–100 km upstream of the mouth of the Altamaha River draining 'undeveloped' areas.

Within North America, the nitrate content of the Mississippi River (Fig. 3) has at least doubled to $\sim$150 µg-atom NO_3 l^{-1} during spring flood over the past 25 yr, while 30 µg-atom NO_3 l^{-1} can still be found downstream on the otherwise oligotrophic shelf at its mouth[15]. Data[46–49] for US rivers which drain other heavily populated areas suggest that their nitrogen content has also increased over the past 25–50 yr from sewage inputs, agricultural fertilizers and nitrate release to groundwater after deforestation[50]. Similarly, US rivers which drain areas of low populations used to contain 10-fold less inorganic nitrogen ($\sim$7–10 µg-atom N l^{-1}) than the above rivers before 1970, but the nitrate content of the Altamaha (Fig. 4) and Pamlico[52] rivers is now $\geq$30 µg-atom NO_3 l^{-1}.

The location and discharge of the world's 30 largest rivers further suggests that a similar freshwater flux occurs from 'undeveloped' areas ($\sim 286\times 10^3$ m^3 s^{-1}) compared with those from 'developed' areas ($\sim 234\times 10^3$ m^3 s^{-1}) with a sixfold larger human population, that is greater anthropogenic nitrogen sources. If we assume that half the world's total river discharge ($\sim 4\times 10^{16}$ l yr^{-1}) has not yet been affected by deforestation, sewage and fertilizers in these underdeveloped countries, and that the low nitrate contents of the Amazon[53], Congo[54], Orinoco (J. Edmond, personal communication) and Niger[54] Rivers ($\leq$10 µg-atom NO_3 l^{-1}) are typical of these regions (Table 3), nitrate input to the shelves from such drainage systems is $\sim 0.3\times 10^{13}$ g NO_3–N yr^{-1}. If we also assume that the other half of the world's river discharge now has at least 60 µg-atom NO_3 l^{-1} due to population distribution and the industrial revolution (>70 µg-atom NO_3 l^{-1} is found in the Hudson, Potomac, Susquehanna, Delaware, Mekong[55] and Yantze (J. Edmond, personal communication) rivers and recall concentrations of 150–300 µg-atom NO_3 l^{-1} in the Mississippi, Seine, Rhine and Scheldt rivers) the annual nitrate input from these developed terrestrial ecosystems would be $\sim 1.7\times 10^{13}$ g NO_3–N yr^{-1}.

A present input of 2.0×10^{13} g NO_3–N yr^{-1} from river runoff would be at least treble that of 0.6×10^{13} g NO_3–N yr^{-1} estimated more than 25 yr ago[56], from data taken more than 50 yr ago[57]. Using a similar unrevised estimate of 0.8×10^{13} g NO_3–N yr^{-1} in river runoff, it has been suggested[58] that only $\sim$10% of the nitrogen fertilizers previously applied to fields had yet to be released from the soil to streams by 1970. In terms of a seasonal range, the inorganic nitrogen content of rural land runoff in Ohio[59] varies during the year from $\sim$10 to 600 µg-atomic N l^{-1} as reflected in the monthly nitrate content of the downstream Mississippi River (Fig. 3). The long-term trend of this and the Altamaha (Fig. 4) time series over the past 25 yr, however, suggests that NO_3 concentrations of these rivers has doubled within the past decade. If mobilization of nitrogen from soil pools is now increasing, we would expect future increases of primary production in the coastal zone and greater export of this carbon to the slope.

In fact, other estimates[60] of riverine nitrogen loading range as high as 3.5×10^{13} g N yr^{-1} and rivers are not the only source of anthropogenic nitrogen inputs to coastal waters. An analysis[61] of the present annual nitrogen loading from coastal sewage outfalls, waste dumping and river input to the shelves of the New York Bight and the Mediterranean, Seto Inland, Baltic, North and Irish Seas suggest a range of 1–3×10^4 tons N yr^{-1} (10^6 persons)$^{-1}$ for each of these areas. Extrapolating the mean of this estimate to half of the world population of 30×10^8 persons for just coastal communities of North America–Asia–Europe (Table 3), we obtain an estimate of 3×10^7 tons N yr^{-1}, or 3×10^{13} g N yr^{-1} from both rivers and other sources of nitrogen within the developed regions.

Conclusions

With a C/N ratio of 5:1 for primary production/nitrogen uptake, this anthropogenic loading of 'new' nitrogen would allow a coastal export of 0.15×10^9 tons C yr^{-1} without considering the carbon fixed in recycling this nitrogen source. The dissolved nitrogen pool of nearshore shallow waters is recycled $\sim$10 times during the year[62], which is reflected by observations that indicate recycled nitrogen amounts to 90% of the daily ^{15}N uptake of phytoplankton in the Hudson[63] and Chesapeake[64] plumes. These data suggest a primary production of 1.50×10^9 tons C yr^{-1} from this source, that is, $\sim$25% of the present shelf production. After each change of the particulate C/N ratio from 5/1 to 10/1 during the nearshore remineralization cycle, 0.75×10^9 tons C of refractory carbon could be available for export each year as a result of this nitrogen loading. Once produced on the outer and inner shelf, organic matter of a C/N ratio of either 5/1 or 10/1 tend to be transported with those fine-grain sediment particles, which, escaping the sediment traps of the estuaries, transit the continental shelf and are deposited on the upper slopes[65,66].

Our analyses suggest that up to half of the estimated 1.50×10^9 tons C yr^{-1} of the slope carbon sink could result from increased eutrophication since the industrial revolution. A combination of outer shelf ungrazed phytodetritus with a C/N of 5/1 and nearshore reworked detritus with a C/N of 10:1, sequestered together on the slope, would constitute both the missing carbon (Tables 1, 2) and the particle flux required as a sink for ^{210}Pb budgets[67]. The biological export of shelf carbon in transient and steady-state budgets of the global CO_2 cycle could present an important management strategy. Increased eutrophication of a well-flushed coastal zone should lead to both increased primary production (fixation of CO_2) and export, not anoxia[68], such that sewage waste disposal problems might be coupled with averting increases in the CO_2 content of the atmosphere.

This research was mainly sponsored by the Department of Energy (DOE) under contract DE-AC02-76CH00016 as part of our Atlantic Coastal Ecosystem (ACE) programme. Additional support was provided by the NSF as part of the Processes and Resources of the Bering Sea (PROBES) programme. Dr Charles Demas, Melvin Edwards and Dr John Edmond provided data on nitrogen content of the Mississippi, Altamaha, Yantze and Orinoco Rivers, while our colleagues in the ACE and PROBES programmes shared additional data and ideas.

Received 8 December 1980; accepted 5 March 1981.

1. Walsh, J. J. in *Primary Productivity in the Sea* (ed. Falkowski, P. G.) 497 (Plenum, New York, 1980).
2. Chamberlin, T. C. *J. Geol.* **7,** 575 (1899).
3. Broecker, W. S., Takahashi, T., Simpson, H. J. & Peng, T. H. *Science* **206,** 409 (1979).
4. Machta, L. & Telegados, K. in *Weather and Climate Modification* (ed. Hess, W. N.) 687 (Wiley, New York, 1974).
5. Machta, L., Hanson, K. & Keeling, C. D. in *The Fate of Fossil Fuel CO_2 in the Oceans* (eds Andersen, N. R. & Malahoff, A.) 131 (Plenum, New York, 1977).
6. Bolin, B. *Science* **196,** 613 (1977).
7. Hoffert, M. I., Callegari, A. J. & Hisieh, C. T. in *Carbon Cycle Modeling* (ed. Bolin, B.) (Wiley, New York, in the press).
8. Broecker, W. S. & Takahashi, T. in *The Fate of Fossil Fuel CO_2 in the Oceans* (eds Andersen, N. R. & Malahoff, A.) 213 (Plenum, New York, 1977).
9. Delcourt, H. R. & Harris, W. F. *Science* **210,** 321 (1980).
10. Gentry, A. H. & Lopez-Parodi, J. *Science* **210,** 1356 (1980).
11. Woodwell, G. M. *et al. Science* **199,** 141 (1978).
12. Ralston, C. W. *Science* **204,** 1345 (1979).
13. Richey, J. E., Brock, J. T., Naiman, R. J., Wissmar, R. C. & Stallard, R. F. *Science* **207,** 1348 (1980).
14. Walsh, J. J., Premuzic, E. T. & Whitledge, T. E. in *Ecohydrodynamics* (ed. Nihoul, J. C. J.) (Elsevier, Amsterdam, in the press).
15. Deuser, W. G. *Org. Geochem.* **1,** 243 (1979).
16. Eppley, R. W. & Petersen, B. J. *Nature* **282,** 677 (1980).
17. Honjo, S. *J. mar. Res.* **38,** 53 (1980).
18. Walsh, J. J. *Limnol. Oceanogr.* **21,** 1 (1976).
19. Yentsch, C. M., Yentsch, C. S. & Strube, L. R. *J. mar. Res.* **35,** 537 (1977).
20. MacIsaac, J. J. & Dugdale, R. C. *Deep-Sea Res.* **19,** 209 (1972).
21. Eppley, R. W., Renger, E. H. & Harrison, W. G. *Limnol. Oceanogr.* **24,** 483 (1979).
22. Conway, H. L. & Whitledge, T. E. *J. mar. Res.* **37,** 657 (1979).
23. Olson, R. J. *Limnol. Oceanogr.* **25,** 1064 (1980).
24. Coachman, L. K. & Walsh, J. J. *Deep-Séa Res* (in the press).
25. Riley, G. A. *Bull. Bingham Oceanogr. Coll.* **19,** 72 (1967).
26. Walsh, J. J. *Deep-Sea Res.* **22,** 201 (1975).
27. Walsh, J. J. *et al. Limnol. Oceanogr.* **23,** 659 (1978).
28. Walsh, J. J. *Nature* (in the press).
29. Dugdale, R. C. & Goering, J. J. *Limnol. Oceanogr.* **12,** 196 (1967).
30. Walsh, J. J. in *Analysis of Marine Ecosystems* (ed. Longhurst, A. R.) 159 (Academic, New York, 1980).
31. Mills, E. L. & Fournier, R. O. *Mar. Biol.* **54,** 101 (1979).
32. Houghton, R. W., Smith, P. C. & Fournier, R. O. *J. Fish. Res. Bd Can.* **35,** 414 (1978).
33. Wong, C. S. *Science* **200,** 197 (1978).
34. Revelle, R. & Suess, H. E. *Tellus* **9,** 18 (1957).
35. Stuiver, M. *Science* **199,** 253 (1978).
36. Keeling, C. D. *Science* **202,** 1109 (1978).
37. Siegenthaler, U. & Oeschger, H. *Science* **199,** 288 (1978).
38. Broecker, W. S. *J. geophys. Res.* **84,** 3218 (1979).
39. Walsh, J. J. *Rapp. P.-V. Reun. Cons. Int. Explor. Mer* **173,** 182 (1978).
40. Fonselius, S. H. *Tellus* **22,** 533 (1970).
41. Postma, H. *Rapp. R.-V. Reun. Cons. Int. Explor. Mer* **172,** 350 (1978).
42. Postma, H. in *North Sea Science* (ed. Goldberg, E. D.) 326 (MIT, Cambridge, 1973).
43. Van der Eijk, M., in *Marine Production Mechanisms* (ed. Dunbar, M. J.) 197 (Cambridge University Press, 1979).
44. Johnston, R. in *North Sea Science* (ed. Goldberg, E. D.) 293 (MIT, Cambridge, 1973).
45. Mommaerts, J. P., Baeyens, W. & Decadt, G. in *Actions de Recherche concertees*, 215 (Bruxelles, 1979).
46. Wolman, M. G. *Science* **174,** 905 (1971).
47. Carpenter, J. H., Pritchard, D. W. & Whaley, R. C. in *Eutrophication: Causes, Consequences, Correctives*, 210 (National Academy of Sciences, Washington, 1969).
48. Kiry, P. R. *Contr. Dept. Limnol. Acad. nat. Sci. Phil.* **4,** 1 (1974).
49. Deck, B. L. thesis, Columbia Univ. (1981).
50. Likens, G. E., Borman, F. H., Pierce, R. S. & Reinners, W. A. *Science* **199,** 492 (1978).
51. Haines, E. B. in *Estuarine Research* Vol. 1 (ed. Cronin, L. E.) 303 (Academic, New York, 1975).
52. Hobbie, J. E., Copeland, B. J. & Harrison, W. G. in *Estuarine Research* Vol. I (ed. Cronin, L. E.) 287 (Academic, New York, 1975).
53. Ryther, J. H., Menzel, D. W. & Corwin, N. *J. mar. Res.* **25,** 69 (1967).
54. Van Bennekom, A. J., Berger, G. W., Helder, W. & deVries, R. T. P. *Neth. J. Sea. Res.* **12,** 296 (1978).
55. Meybeck, M. & Carbonnel, J. P. *Nature* **255,** 134 (1975).
56. Emery, K. O., Orr, W. L. & Rittenberg, S. C. in *Essays in Honor of Captain Allan Hancock*, 299 (University of South California Press, Los Angeles, 1955).
57. Clarke, F. W. *Bull. U.S. geol. Surv.* **770,** 1 (1924).
58. Garrels, R. M., MacKenzie, F. T. & Hunt, C. *Chemical Cycles and the Global Environment*, 1 (Kaufmann, Los Altos, 1973).
59. Weibel, S. R., Weidner, R. B., Cohen, J. M. & Christianson, A. G. *J. Am. Wat. Works Ass.* **58,** 1075 (1966).
60. Delwiche, C. C. & Likens, G. E. in *Global Chemical Cycles and their Alteration by Man* (ed. Stumm, W.) 89 (Dahlem Konferenzen, Berlin, 1977).
61. Segar, D. *MESA Spec. Rep.* Ser. (in the press).
62. Smayda, T. J. in *Effects of Energy-related Activities on the Atlantic Continental Shelf* (ed. Manowitz, B.) 70 (BNL 50484, 1976).
63. Garside, C. *Limnol. Oceanogr.* (in the press).
64. McCarthy, J. J., Taylor, W. R. & Taft, J. L. *Limnol. Oceanogr.* **22,** 996 (1977).
65. Schubel, J. & Okubo, A. in *Shelf Sediment Transport: Process and Pattern* (eds Swift, D. J., Duane, D. B. & Pilkey, O. H.) 33 (Dowden, Hutchinson & Ross, Stroudsburg, 1972).
66. Hay, W. W. & Southam, J. R. in *The Fate of Fossil Fuel CO_2 in the Oceans*, (eds Anderson, N. R. & Malahoff, A.) 569 (Academic, New York, 1977).
67. Spencer, D. W., Bacon, M. F. & Brewer, P. G. *J. mar. Res.* (in the press).
68. Falkowski, P. G., Hopkins, T. S. & Walsh, J. J. *J. mar. Res.* **38,** 479 (1980).

Homology between the primary structure of α-fetoprotein, deduced from a complete cDNA sequence, and serum albumin

Simon W. Law* **& Achilles Dugaiczyk**†

Department of Cell Biology, Baylor College of Medicine, Texas Medical Center, Houston, Texas 77030, USA

The entire DNA complementary to murine α-fetoprotein mRNA has been cloned in the plasmid vector pBR322 and its complete nucleotide sequence determined. The deduced amino acid sequence identifies a hydrophobic prepeptide of 19 amino acids and the complete primary structure of α-fetoprotein. Its structure reveals extensive homology to serum albumin and suggests that there are 15 disulphide bridges in the α-fetoprotein molecule, all at the same positions as those found in albumin. The AFP molecule can be organized into a three-domain structure almost identical to that of serum albumin, suggesting a common ancestral origin of the two proteins.

α-FETROPROTEIN (AFP), a serum protein with an apparent molecular weight[1] of 72,000, is made and secreted by the fetal liver[2] and yolk sac[3] of all mammals so far studied. The serum concentration of AFP decreases rapidly during fetal development to a residual level in adulthood but rises again in individuals with hepatomas or germ-cell tumours[4–6]. Elevation of serum AFP levels is also observed in a variety of pathological conditions[6,7] and in certain human fetal malformations such as neural tube defects[8,9]. The exact function of AFP is unclear, although similar biological and physicochemical properties suggest that it may be the fetal counterpart of adult serum albumin.

Human AFP and albumin have been reported to exhibit some homology in amino acid sequence[10]. However, a comparison of the amino-terminal sequence of human, murine and rat AFP with sequences of the respective albumins showed no sequence homology between the two proteins[11]. Furthermore, cDNA to rat AFP mRNA does not cross-hybridize to rat albumin mRNA in moderately stringent conditions[12,13]. The determination of a partial cDNA sequence of rat AFP[14,15] has recently revealed a partial homology between the rat AFP and serum albumin

---

* Present address: National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda, Maryland 20205, USA.

† To whom correspondence should be addressed.

0028-0836/81/210201—05$01.00

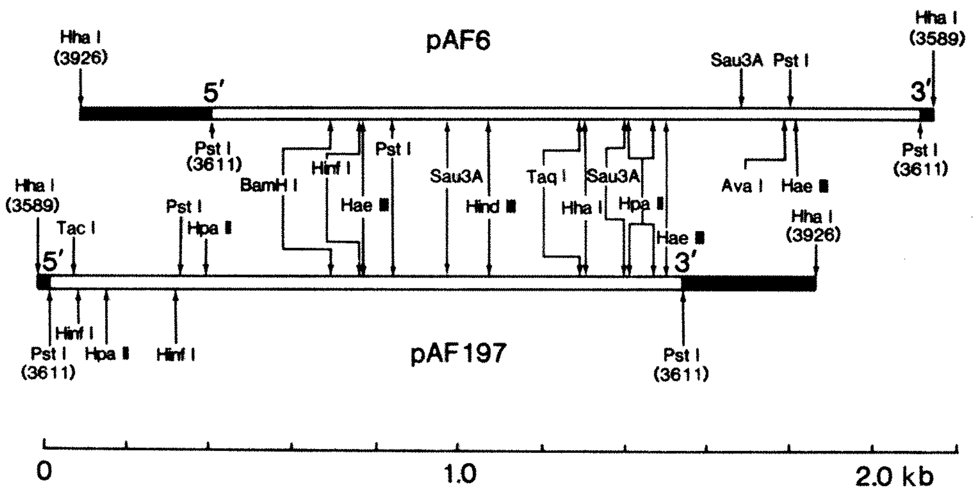


Fig. 1. Restriction endonuclease cleavage sites on the overlapping murine AFP cDNA inserts in the recombinant plasmids pAF6 and pAF197. The map shows the inserted AFP DNA (open bar) and its orientation in the PBR322 plasmid DNA (black bar). The murine-derived DNA is inserted at the *Pst*I site of the plasmid DNA as described in detail previously[16]. The AFP mRNA sequence is presented by the upper DNA strand of the pAF6 and pAF197 map, with the 5′ and 3′ ends as indicated. Numbers on restriction sites within the pBR322 sequence, like *Pst*I (3,611), are taken from available sequence data from Sutcliffe[35] and represent positions on the sequence map of pBR322, counted from the single *Eco*RI site on the circular DNA molecule. kb, Kilobases.

amino acid sequence. We now describe the cloning and nucleotide sequence determination of the entire DNA complementary to murine AFP mRNA and analyse the deduced primary amino acid structure and disulphide bonding pattern of the AFP molecule in the light of its homology to serum albumin.

The recombinant plasmid pAF6

We have previously reported[16] the construction of a cDNA clone containing about 1.65 kilobases of mouse AFP mRNA sequence (Fig. 1). The clone was identified on the basis of a hybrid-arrested *in vitro* translation of mouse yolk sac mRNA, which specifically arrested the translation of the AFP mRNA species. As shown below, nucleotide sequence analysis has revealed that the cDNA sequence in pAF6 corresponds to the 3′ region of the AFP message. This clone does not contain the 5′ sequence corresponding to the known 20 amino acids at the amino terminus of mouse AFP. Comparison of the size of the protein[1] and the size of the cDNA in pAF6 indicates that about 400 nucleotides of additional coding sequence are required to complete the 5′-terminal sequence of the AFP message.

Cloning of cDNA complementary to the 5′ region of AFP mRNA

A *Bam*HI–*Ava*I DNA fragment was isolated from the cDNA clone pAF6 (Fig. 1), annealed to mouse yolk sac mRNA and the resulting RNA/DNA template/primer used in the reverse transcriptase reaction to synthesize a cDNA copy for the 5′ region of the mRNA template. Using the same reverse transcriptase the cDNA was converted to its double-stranded form[16], treated with endonuclease S_1, tailed with about 15 (dC)-residues and annealed with the *Pst*I-linearized plasmid vector pBR322 which had been tailed with 15 (dG)-residues as previously described[17]. Clones of the *Escherichia coli* strain RR1[18] which had been transformed[19] by the annealed recombinant plasmids were identified by their tetracycline resistance and ampicillin sensitivity, the two markers being carried on the plasmid pBR322 genome[18].

Clones containing sequences 5′ upstream of the *Bam*HI–*Ava*I DNA primer were identified by colony hybridization[20]; the probe for this hybridization was obtained by using polynucleotide kinase and [γ-^{32}P]ATP to end label[21] the 5′-terminal *Pst*I–*Bam*HIDNA fragment of pAF6 (Fig. 1), containing the AFP cDNA sequence 5′ upstream of the *Bam*HI–*Ava*I DNA primer. Sequence analysis of one of the positive clones, pAF197, which was found to contain about 400 base pairs 5′ upstream of the last cDNA nucleotide of pAF6 (Fig. 1), revealed that this cDNA codes for the entire amino-terminal region of murine AFP up to amino acid −9 of the AFP prepeptide; a partial amino acid sequence of the prepeptide has recently been reported[22]. As the two cDNA clones share an extensive (1.1 kilobase) segment of homologous DNA in their central portion and pAF6 codes for the 3′-terminal region of AFP mRNA, together they encode the entire sequence of the murine AFP molecule.

DNA sequence analysis of the AFP cDNA

In contrast to serum albumin, where the entire amino acid sequence is known for the human and bovine proteins, very little sequence information is available for AFP, apparently because of hindrance by the carbohydrate moieties of the molecule[11]. What little sequence information is available includes 20 amino acids of the prepetide[22], some 20 residues at the amino terminus and the carboxy-terminal valine[11].

We obtained the amino acid sequence of AFP by determining the entire sequence of the AFP cDNA insert in pAF6 and pAF197 by the chemical method of Maxam and Gilbert[23], mainly using both DNA strands to ensure accuracy. All the restriction sites on the restriction map were accounted for by sequence analysis, and all the restriction sites used to end-label DNA fragments were sequenced across by labelling a neighbouring restriction site. The murine-derived sequence of pAF197 actually starts with the *Mnl*I recognition sequence CCTC and corresponds to leucine −9 of the prepeptide. The sequence 5′ upstream of leucine −9 was determined as described in Fig. 2 legend.

Figure 2 shows entire nucleotide sequence of the cDNA in pAF6, pAF197, of the primer-extended cDNA at the 5′ terminus of the AFP mRNA and of the deduced amino acid sequence. The combined length is 2,064 nucleotides, of which 41 represent the 5′-untranslated region, 57 identify the prepeptide of 19 amino acids, 1,758 code for the 586 amino acids of AFP, 153 make up the 3′-untranslated region and 55 are the poly(A) sequence.

The hydrophobic presequence (Fig. 2) for AFP, which is 19 amino acids long, is processed to generate the mature polypeptide. The length and hydrophobic nature of this presequence are similar to those reported for other secretory proteins (see ref. 24 for a compilation of presequences). Instead of two terminal methionine residues[22] only one is accountable for by direct DNA sequencing. No other ATG codons are observed in any reading frame in the entire DNA sequence 5′ upstream from the amino-terminal methionine. Therefore, our data on the presequence show a length of 19 rather than 20 amino acids as previously reported[22]. The remaining DNA sequence data agree, with few exceptions, with the published amino acid sequence in this region of AFP. The amino-terminal residue in AFP is found to be Ala rather than the previously reported Ser[11]. Considering that our sequence in this region is derived

```
                                          -19                                 -10                                 -1
                                          met lys trp ile thr pro ala ser leu ile leu leu leu his phe ala ala ser lys
TCCCACTTCCAGCACTGCCTGCGGTGAAGGAACAAGCAGCC ATG AAG TGG ATC ACA CCC GCT TCC CTC ATC CTC CTG CTA CAT TTC GCT GCG TCC AAA (98)

1                                   10                              18      20                                      30
ala leu his glu asn glu phe gly ile ala ser thr leu asp ser ser gln cys val thr glu lys asn val leu ser ile ala thr ile
GCA TTG CAC GAA AAT GAG TTT GGG ATA GCT TCC ACG TTA GAT TCC TCC CAG TGC GTG ACG GAG AAG AAT GTG CTT AGC ATA GCT ACC ATC (188)

31                                  40                                      50                                      60
thr phe thr gln phe val pro glu ala thr glu glu glu val asn lys met thr ser asp val leu ala ala met lys lys asn ser gly
ACC TTT ACC CAG TTT GTT CCG GAA GCC ACC GAG GAG GAA GTG AAC AAA ATG ACT AGC GAT GTG TTG GCT GCA ATG AAG AAA AAC TCT GGC (278)

61      63                          70                      76              80                                      90
asp gly cys leu glu ser gln leu ser val phe leu asp glu ile cys his glu thr glu leu ser asn lys tyr gly leu ser gly cys
GAT GGG TGT TTA GAA AGC CAG CTA TCT GTG TTT CTG GAT GAA ATT TGC CAT GAG ACG GAA CTC TCT AAC AAG TAT GGA CTC TCA GGC TGC (368)

91                                  100 101                                 110                                     120
cys ser gln ser gly val glu arg his gln cys leu leu ala arg lys lys thr ala pro ala ser val pro pro phe gln phe pro glu
TGC AGC CAA AGT GGA GTG GAA AGA CAT CAG TGT CTG CTG GCA CGC AAG AAG ACT GCT CCG GCC TCT GTC CCA CCC TTC CAG TTT CCA GAA (458)

121             125                 130                                     140                                     150
pro ala glu ser cys lys ala his glu glu asn arg ala val phe met asn arg phe ile tyr glu val ser arg arg asn pro phe met
CCT GCC GAG AGT TGC AAA GCA CAT GAA GAA AAC AGG GCA GTG TTC ATG AAC AGG TTC ATC TAT GAA GTG TCA AGG AGG AAC CCC TTC ATG (548)

151                                 160                             169 170                                 178     180
tyr ala pro ala ile leu ser leu ala ala gln tyr asp lys val val leu ala cys cys lys ala asp asn lys glu glu cys phe gln
TAT GCC CCA GCC ATT CTG TCC TTG GCT GCT CAG TAC GAC AAG GTC GTT CTG GCA TGC TGC AAA GCT GAC AAC AAG GAG GAG TGC TTC CAG (638)

181                                 190                                     200 201                                 210
thr lys arg ala ser ile ala lys glu leu arg glu gly ser met leu asn glu his val cys ser val ile arg lys phe gly ser arg
ACA AAG AGA GCA TCC ATT GCA AAG GAA TTA AGA GAA GGA AGC ATG TTA AAT GAG CAT GTA TGT TCA GTG ATA AGA AAA TTT GGA TCC CGA (728)

211                                 220                                     230                                     240
asn leu gln ala thr thr ile ile lys leu ser gln lys leu thr glu ala asn phe thr glu ile gln lys leu ala leu asp val ala
AAC CTC CAG GCA ACA ACC ATT ATT AAG CTA AGT CAA AAG TTA ACT GAA GCA AAT TTT ACT GAG ATT CAG AAG CTG GCC CTG GAT GTG GCT (818)

241                 246 247                         254                     260                     266             270
his ile his glu glu cys cys gln gly asn ser leu glu cys leu gln asp gly glu lys val met thr tyr ile cys ser gln gln asn
CAC ATC CAC GAG GAG TGT TGC CAA GGA AAC TCG CTG GAG TGT CTG CAG GAT GGG GAA AAA GTC ATG ACA TAT ATA TGT TCT CAA CAA AAT (908)

271                             279 280                                     290                                     300
ile leu ser ser lys ile ala glu cys cys lys leu pro met ile gln leu gly phe cys ile ile his ala glu asn gly val lys pro
ATT CTG TCA AGC AAA ATA GCA GAG TGC TGC AAA TTA CCC ATG ATC CAA CTA GGC TTC TGC ATA ATT CAC GCA GAG AAT GGC GTC AAA CCT (998)

301                                 310                                     320                                     330
glu gly leu ser leu asn pro ser gln phe leu gly asp arg asn phe ala gln phe ser ser glu glu lys ile met phe met ala ser
GAA GGC TTA TCT CTA AAT CCA AGC CAG TTT TTG GGA GAC AGA AAT TTT GCC CAA TTT TCT TCA GAG GAA AAA ATC ATG TTC ATG GCA AGC(1088)

331                                 340                                     350                                     360
phe leu his glu tyr ser arg thr his pro asn leu pro val ser val ile leu arg ile ala lys thr tyr gln glu ile leu glu lys
TTT CTT CAT GAA TAC TCA AGA ACT CAC CCC AAC CTT CCT GTC TCA GTC ATT CTA AGA ATT GCT AAA ACG TAC CAG GAA ATA TTG GAG AAG(1178)

361                                 370                                     380                                     390
cys ser gln ser gly asn leu pro gly cys gln asp asn leu glu glu glu leu gln lys his ile glu glu ser gln ala leu ser lys
TGT TCC CAG TCT GGA AAT CTA CCT GGA TGT CAG GAC AAT CTG GAA GAA GAA TTG CAG AAA CAC ATC GAG GAG AGC CAG GCA CTG TCC AAG(1268)

391     393                         400                                     410                                     420
gln ser cys ala leu tyr gln thr leu gly asp tyr lys leu gln asn leu phe leu ile gly tyr thr arg lys\ala pro gln leu thr
CAA AGC TGC GCT CTC TAC CAG ACC TTA GGA GAC TAC AAA TTA CAA AAT CTG TTC CTT ATT GGT TAC ACG AGG AAA GCC CCT CAG CTG ACC(1358)

421                                 430                             438 439 440                                 449 450
ser ala glu leu ile asp leu thr gly lys met val ser ile ala ser thr cys cys gln leu ser glu glu lys trp ser gly cys gly
TCA GCA GAG CTG ATC GAC CTC ACC GGG AAG ATG GTG AGC ATT GCC TCC ACG TGC TGC CAG CTC AGC GAG GAG AAA TGG TCC GGC TGT GGT(1448)

451                                 460     462                             470                         477 478     480
glu gly met ala asp ile phe ile gly his leu cys ile arg asn glu ala ser pro val asn ser gly ile ser his cys cys asn ser
GAG GGA ATG GCC GAC ATT TTC ATT GGA CAT TTG TGT ATA AGG AAT GAA GCA AGC CCT GTG AAC TCT GGT ATC AGC CAC TGC TGC AAC TCT(1538)

481                         488     490                                     500                                     510
ser tyr ser asn arg arg leu cys ile thr ser phe leu arg asp glu thr tyr ala pro pro pro phe ser glu asp lys phe ile phe
TCG TAT TCC AAC AGG AGG CTA TGC ATC ACC AGT TTT CTG AGG GAT GAA ACC TAT GCC CCT CCC CCA TTC TCT GAG GAT AAA TTC ATC TTC(1628)

511             514                 520                                     530                                     540
his lys asp leu cys gln ala gln gly lys ala leu gln thr met lys gln glu leu leu ile asn leu val lys gln lys pro glu leu
CAC AAG GAT CTG TGC CAA GCT CAG GGC AAA GCC CTA CAG ACC ATG AAA CAA GAG CTT CTC ATT AAC CTG GTG AAG CAA AAG CCT GAA CTG(1718)

541                                                                 559 560                                 568     570
thr glu glu gln leu ala ala val thr ala asp phe ser gly leu leu glu lys cys cys lys ala gln asp gln glu val cys phe thr
ACA GAG GAG CAG CTG GCG GCT GTC ACT GCA GAT TTC TCG GGC CTT TTG GAG AAG TGC TGC AAA GCC CAG GAC CAG GAA GTC TGT TTC ACA(1808)

571                                 580
glu glu gly pro lys leu ile ser glu thr arg asp ala leu gly val ter                                     ter
GAA GAG GGT CCA AAG TTG ATT TCC GAA ACT CGT GAT GCT TTG GGC GTT TAA ACATCTCCAGAAGGAAGAGTGGACAAAAAAATGTGTTGACGCTTTGGTGTG(1910)


            ter   ter            ter
AGCCTTTTGGCTTAACTGTAACTGCTAGTACTTTAACCACATGGTGAAGATGTCCATGTGAGATTTCTATACCTTAGGAATAAAAACTTTTCAACTATTAAA..... 50 ......AA(2064)
```

Fig. 2 Nucleotide sequence of murine AFP mRNA as determined from cloned cDNA. The Maxam and Gilbert[23] method was used throughout this work for sequence determination. The sequence 5′ upstream of leucine −9, which is not contained in pAF197, was determined as follows. A 0.1-kilobase *Hha*I(3,589)–*Hin*fI DNA fragment (Fig. 1) was isolated from pAF197 and 5′ labelled with polynucleotide kinase[21] and [γ-^{32}P]ATP. This fragment was subsequently digested with *Mnl*I, which cuts at the very beginning of the AFP sequence and removes the pBR322 portion of the *Hha*I–*Hin*fI fragment. The resulting *Mnl*I–*Hin*fI murine DNA fragment was labelled only at the *Hin*fI 5′ terminus. This fragment was annealed to mouse yolk sac mRNA and the resulting RNA/DNA template/primer was used in the reverse transcriptase reaction to synthesize a cDNA copy extended into the 5′ terminus of the mRNA template. This cDNA was separated from the deoxyribonucleotide triphosphates on a Sephadex G-50 column and used directly for sequence determination. The amino acid sequence shown is deduced from the nucleotide sequence. Nucleotides are numbered at the end of each row and amino acids above the sequence. In the 3′-untranslated region four potential termination codons are indicated in the amino acid reading frame. There seems to be no propeptide sequence in AFP analogous to that in serum albumin. The molecular weights for the aglycoprotein of AFP and its precursor, as determined from the amino acid composition, are 65,164 and 67,286, respectively.

from sequencing both DNA strands, we reassign the amino acid positions accordingly (Fig. 2).

The 5′-noncoding region of the murine AFP message reveals the presence of a purine-rich sequence in positions 26–38 (Fig. 2). This region exhibits partial sequence homology among the various mammalian mRNAs which have been found to have a purine-rich region with its 3′ end 2–6 nucleotides from the starting AUG codon (for a compilation of 5′-noncoding sequences see refs 25, 26). Nucleotides 6–14 (Fig. 2) are complementary to a 3′-terminal region of eukaryotic 18S

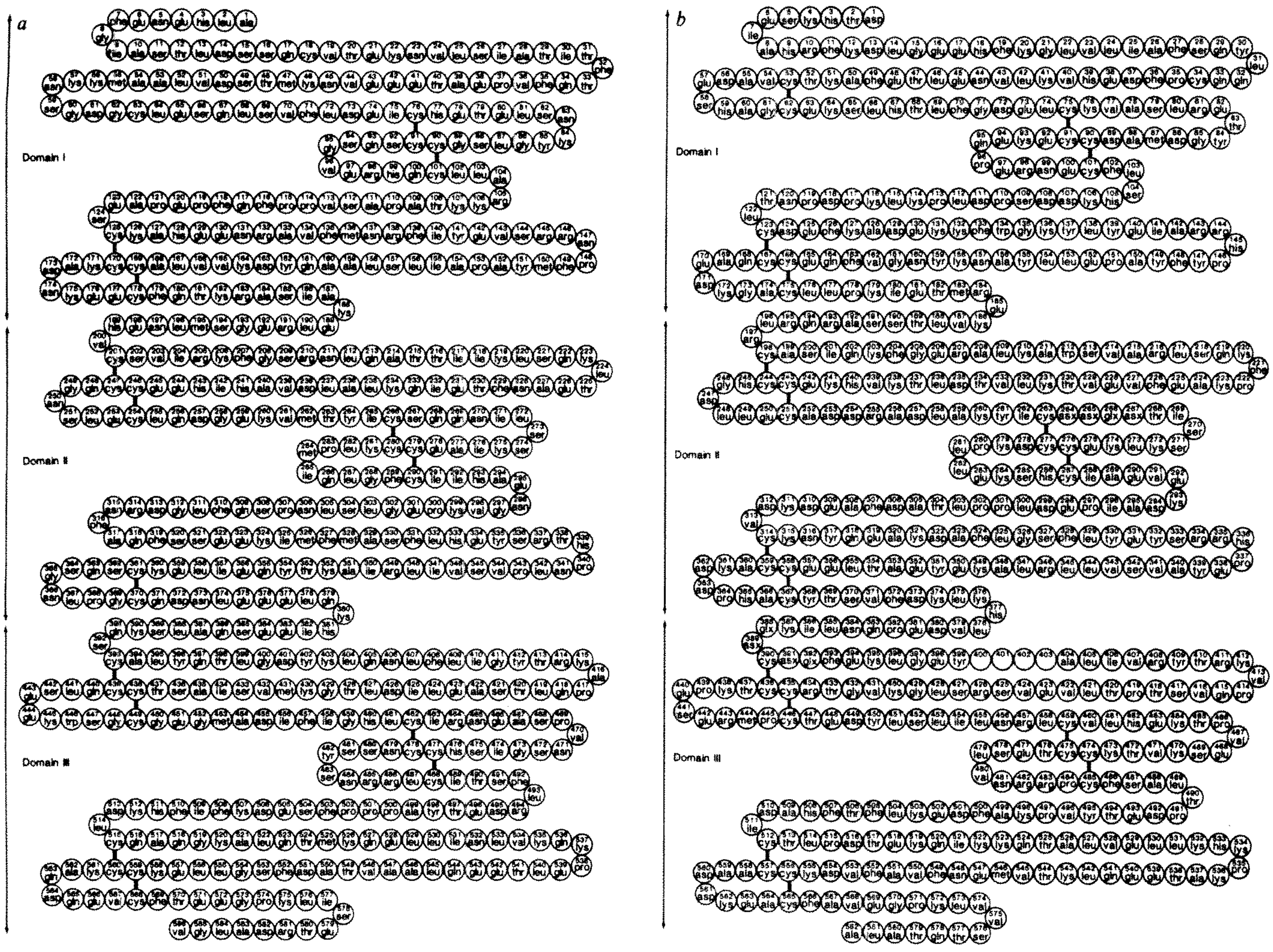


Fig. 3 *a*, Amino acid sequence and disulphide bonding pattern of murine AFP. A triple domain structure is indicated by arrows: I, amino acids 1–193; II, amino acids 194–385; III, amino acids 386–586. The sequence represents our present data, the layout of the sequence follows that of serum albumin, and the disulphide bridges are assigned by analogy with serum albumin. *b*, Amino acid sequence and disulphide bonding pattern of bovine serum albumin. The triple domain structure is indicated by arrows: I, amino acids 1–190; II, amino acids 191–382; III, amino acids 383–582. The sequence and its layout are according to Brown[33].

RNA[27] and thus could represent a ribosome binding site:

(5′)...CUUCC^GCA..............AFP mRNA

(3′)...GAAGGCGUCCm_2^6A m_2^6A...18S RNA

The 3′-noncoding region consists of 153 nucleotides beyond the carboxy-terminal valine. It starts with the translation stop signal UAA and contains four more potential stop codons in the same reading frame. The hexanucleotide AAUAAA, which is conserved in this region of eukaryotic mRNAs[28], is located in the AFP message 15 nucleotides 5′ upstream from the poly(A). Contrary to reports on some human cDNA sequences[29], mouse AFP codons are not biased towards those ending in one or two specific nucleotides, although there is a bias in the use of certain codons.

Structural relationship between α-fetoprotein and albumin

From a detailed amino acid sequence analysis of bovine (BSA) and human (HSA) serum albumin[30–33], Brown discovered a repeating pattern in the structure[32,33]. On the basis of spacing of the cysteine residues in the complete polypeptide chain and the positions of the disulphide bridges, the albumin sequence can be organized to generate three very similar repeat units, called domains (Fig. 3*b*).

To elucidate the structure of AFP, we have translated the entire AFP mRNA coding sequence into a corresponding amino acid sequence and aligned the disulphide bridges of cysteine residues in the same way as has been described for BSA or HSA[32,33]. The resulting structure reveals that the amino acid sequence of AFP can be divided into three domains of sequence homology, analogous to the domains of albumin (Fig. 3*a*, b). A total 15 disulphide bridges can be formed throughout the AFP molecule, two less than found in BSA. The cysteine residues occur within the polypeptide chain of AFP at almost identical sequence positions to those found in albumin, thus allowing for the formation of the same disulphide bridges in the secondary structure of AFP. Note however, the disulphide bridges in AFP are assigned here by analogy with serum albumin only, and we have no experimental evidence for such arrangement.

A comparison of the AFP amino acid sequence with that of human or bovine albumin shows about 40, 30 and 10% homology of identical amino acids between domains III, II and I, respectively. The homology between the two proteins increases further to ~65% for domain III and ~55% for domain II if the amino acids are compared on the basis of their belonging to one of four groups having a characteristic side chain: hydrophobic (Phe, Leu, Ile, Met, Val, Pro, Ala, Tyr, Trp), polar (Ser, Thr, Gln, Asn, Cys, Gly), basic (His, Lys, Arg) or acidic (Glu, Asp). Among the differences between the two proteins there are three potential N-glycosylation sites (Asn-X-Thr; Asn-X-Ser)[34] in the AFP sequence and none in albumin.

It is tempting to analyse the nucleotide sequence of the AFP message in the two regions where, in contrast to the albumin sequence, there are no disulphide bridges in the AFP structure. One such region is the Cys 53–Cys 62 disulphide pair in BSA

(Fig. 3*b*). In AFP, this structure corresponds to the juxtapositioned Ala 54–Cys 63 amino acid pair (Fig. 3*a*) which obviously lost a bridging capacity. There is, however, a cysteine UGC codon and thus a 'hidden' Cys residue in the critical position of Ala 54, if the reading frame is shifted by one nucleotide in this region (Fig. 2). A difference of one nucleotide is thus responsible for the difference of one extra disulphide bridge in the structure of the two proteins. The other region is the Cys 314–Cys 359 pair in BSA (Fig. 3*b*), which is also absent from the AFP structure, where it is replaced by the Ala 317–Ser 362 pair instead (Fig. 3*a*). Examination of the AFP nucleotide sequence (Fig. 2) reveals again that only single nucleotide changes are required to re-establish the missing Cys 317–Cys 362 structure in the AFP molecule. The mRNA sequence seems to carry a more detailed record of ancestral relationship between AFP and serum albumin than is evident from their existing amino acid homology. When DNA sequences for AFP and albumin from the same species become available, their comparison will probably give a better indication of the genetic relationship between the two proteins.

This work was supported by the Baylor Center for Population Research and Reproductive Biology. We thank Olivia Dennison and Serlina Robinson for technical assistance, Dr John Compton and James Hawkins for critical reading of the manuscript and David Scarff for the artwork.

Received 28 November 1980; accepted 26 February 1981.

1. Rouslahti, E. & Seppala, M. *Int. J. Cancer* **7,** 218–225 (1971).
2. Abelev, G. I. *Adv. Cancer Res.* **14,** 295–358 (1971).
3. Gitlin, D. *et al. Cancer Res.* **32,** 979–982 (1972).
4. Abelev, G. I. *Transplantn Rev.* **20,** 1–37 (1974).
5. Rouslahti, E., Pihko, H. & Seppala, M. *Transplantn Rev.* **20,** 38–60 (1974).
6. Sell, S., Becker, F., Leffert, H. & Waterbe, H. *Cancer Res.* **36,** 4239–4249 (1976).
7. Waterbe, H. *Cancer Res.* **31,** 1192–1194 (1971).
8. Allan, L. D. *Lancet* ii, 522–525 (1973).
9. Brock, D. J. H. & Sutcliffe, R. G. *Lancet* **ii,** 197–199 (1972).
10. Rouslahti, E. & Terry, W. D. *Nature* **260,** 804–805 (1976).
11. Peters, E. H., Nishi, S. & Tamaoki, T. *Biochem. biophys. Res. Commun.* **83,** 75–82 (1978).
12. Innis, M. A. & Miller, D. L. *J. biol. Chem.* **252,** 8469–8475 (1977).
13. Sala-Trepat, J. M. *et al. Biochemsitry* **18,** 2167–2178 (1979).
14. Innis, M. A. & Miller, D. L. *J. biol. Chem.* **255,** 8994–8996 (1980).
15. Liao, W. S. L., Hamilton, R. W. & Taylor, J. M. *J. biol. Chem.* **255,** 8046–8049 (1980).
16. Law, S. W., Tamaoki, T., Kreuzaler, F. & Dugaiczyk, A. *Gene* **10,** 53–61 (1980).
17. Dugaiczyk, A., Robberson, D. L. & Ullrich, A. *Biochemistry* **19,** 5869–5873 (1980).
18. Bolivar, F. *et al. Gene* **2,** 95–113 (1977).
19. Mandel, M. & Higa, A. *J. molec. Biol.* **53,** 159–162 (1970).
20. Grustein, M. & Hogness, D. S. *Proc. natn. Acad. Sci. U.S.A.* **72,** 3961–3965 (1975).
21. Lillehaug; J. R. & Kleppe, K. *Biochemistry* **14,** 1225–1229 (1975).
22. Peters, E. K. *et al. Cancer Res.* **39,** 3702–3706 (1979).
23. Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65,** 499–560 (1980).
24. Habener, J. F. *et al. Proc. natn. Acad. Sci. U.S.A.* **75,** 2616–2620 (1978).
25. Baralle, F. E. & Brownlee, G. G. *Nature* **274,** 84–87 (1978).
26. Kozak, M. *Cell* **15,** 1109–1123 (1978).
27. Azad, A. A. & Deacon, N. J. *Nucleic Acids Res.* **8,** 4365–4376 (1980).
28. Proudfoot, N. J. & Brownlee, G. G. *Nature* **263,** 211–214 (1976).
29. Fiddes, J. C. & Goodman, H. M. *Nature* **286,** 684–687 (1980).
30. Meloun, B., Moravek, L. & Kostka, V. *FEBS Lett* **58,** 134–137 (1975).
31. Brown, J. R. *Fedn Proc.* **34,** 591 (1975).
32. Brown, J. R. *Fedn Proc.* **35,** 2141–2144 (1976).
33. Brown, J. R. *FEBS 11th Meet., Copenhagen* **50,** 1–10 (1977).
34. Spiro, R. G. *Adv. Protein Chem.* **27,** 349–467 (1973).
35. Sutcliffe, J. G. *Nucleic Acids Res.* **5,** 2721–2728 (1978).

Nucleotide sequence of *Xenopus laevis* 18S ribosomal RNA inferred from gene sequence

Mohammad Salim & B. Edward H. Maden

Department of Biochemistry, University of Glasgow, Glasgow G12 8QQ, UK

18S ribosomal RNA in Xenopus laevis *is 1,825 nucleotides long, as inferred from sequence analysis of an 18S gene. All the 40 rRNA methyl groups can be located in the sequence. Comparison with the yeast* (Saccharomyces cerevisiae) *18S sequence reveals extensive regions of high homology interspersed with tracts having little or no homology. Regions of high homology contain almost all the RNA methyl groups. Major regions of low homology are considerably richer in C + G in* Xenopus *than in yeast.*

KNOWLEDGE of the primary structure of ribosomal RNA (rRNA) in higher organisms should elucidate several aspects of ribosomal function, assembly and phylogeny. Here we report the nucleotide sequence of 18S rRNA in the frog *Xenopus laevis*, as inferred from analysis of an 18S gene. We compare our findings with the recently reported yeast (*Saccharomyces cerevisiae*) 18S sequence[1].

The sequence was determined from the *X. laevis* rDNA clone, pXlr 101, which contains a complete ribosomal transcription unit with its preceding non-transcribed spacer (Fig. 1). The sequencing strategy involved almost complete analysis of both strands of the 18S gene with extensive overlapping (Fig. 1); the inferred RNA sequence is shown in Fig. 2. Considerable evidence indicates that this sequence is representative of 18S rRNA in *X. laevis*. Oligonucleotide data on the 5′ 160 nucleotides[2], the 3′ 230 nucleotides (B.E.H.M., unpublished data), the various oligo(A) tracts[3] and the methylated sequences (below) all agree with the DNA sequence. Limited sequence data from other clones of *X. laevis* rDNA agree with the sequence shown in Fig. 2 (B.E.H.M., unpublished results), and the many restriction sites identified here agree well with those reported previously for an independently cloned rDNA unit from *X. laevis*[4,5].

rRNA methylation

Knowledge of the sequence provides a means of determining the sites of all the rRNA methyl groups (40 in *X. laevis* 18S rRNA). Data on the methylated oligonucleotides in 18S RNA[6–10] and on their approximate locations in the sequence as deduced from 'hybridization mapping'[11] have allowed us to locate all the methylated nucleotides. The methyl group distribution is highly irregular (see Figs 2, 3). There is considerable clustering of 2′-*O*-methyl groups in the 5′ 40% of the sequence and the distribution in this region is also irregular; another cluster of 2′-*O*-methyl groups occurs in the region 1,240–1,400. The remaining 2′-*O*-methyl groups are outliers from the main clusters. All the base methyl groups are in the 3′ 35% of the molecule. The hypermodified nucleotide amψ (1-methyl-3-aminobutyryl pseudouridine[12]) is at position 1,209. m^7G occurs at the point encoded by the *Eco*RI site (nucleotide 1,596). The base-modified A residues are in the last 40 nucleotides.

The numerous pseudouridines in 18S rRNA are also found by hybridization mapping to be widely distributed (unpublished results), but, as few have been precisely located[2], this class of modifications is not included in Fig. 2.

0028-0836/81/210205—04$01.00

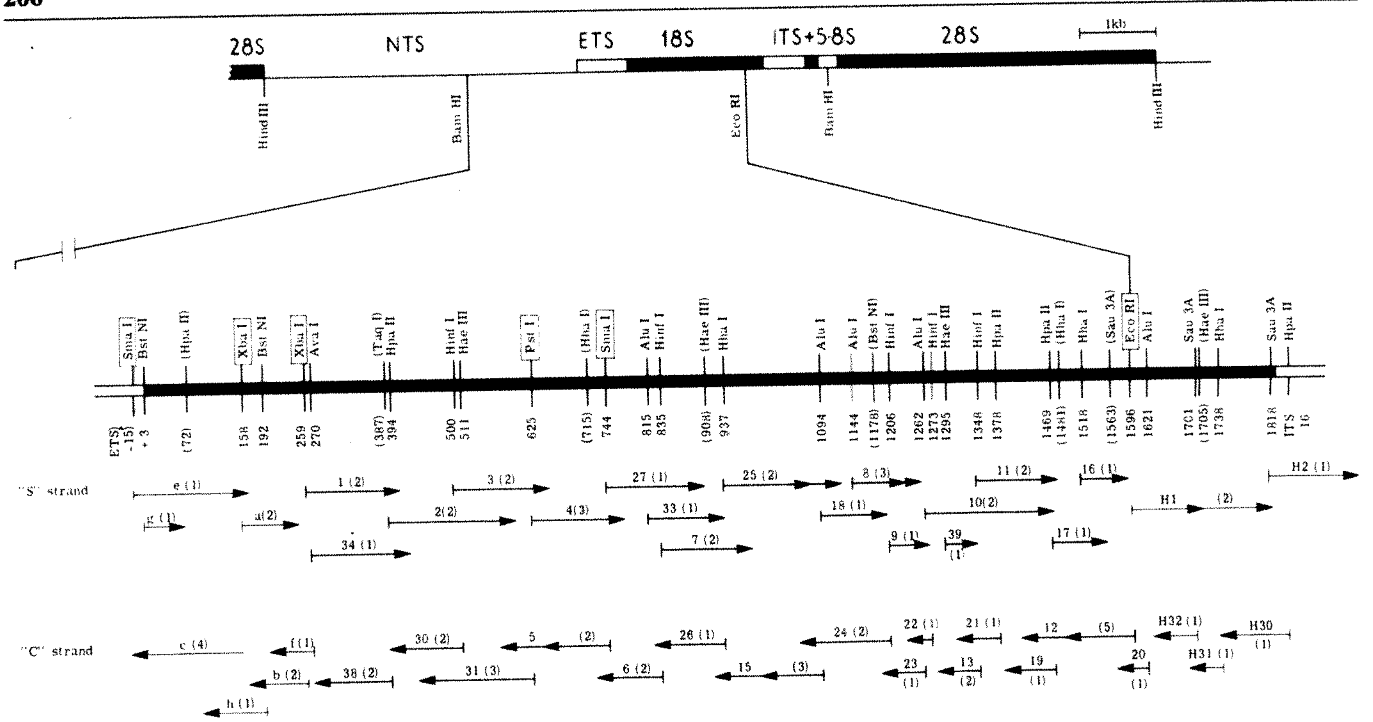


Fig. 1 Top line, repeating structure of rDNA in *X. laevis*. Solid sections represent rRNA coding sequences. NTS, ETS and ITS denote non-transcribed, external and internal transcribed spacers. pX1r 101 consists of a complete rDNA unit bounded by *Hin*dIII sites, cloned in the vector pMB9. Lower part of the figure shows the sequencing strategy. Most of the sequence was determined after subcloning the indicated 3.5-kilobase *Bam*HI/*Eco*RI fragment into pBR322 (ref. 2). The sequence to the right of *Eco*RI sites was determined after subcloning the 1.1-kilobase *Eco*RI/*Bam*HI fragment into pBR322 (ref. 30). The sequence through the *Eco*RI site was determined as indicated below (gel 20). Sequences of the 5′ 275 nucleotides and 3′ 227 nucleotides, including identification of the two ends of the 18S gene, have been reported elsewhere[2,30]; sequencing strategy for these regions is included for completeness. The major part of the sequence was determined by the Maxam and Gilbert method[31] from gels 1–38. Restriction sites used for sequencing are shown on the centre line; numbers denote the first nucleotide in each site counted from the 5′ end of the 18S gene. Boxed sites were used separately or in combination for large-scale digests. Some of the resulting material was labelled at the 5′ termini; the remainder was digested with one or more further enzymes before labelling. Upper arrows denote sequencing gels on the 'S' strand (synonymous to RNA); lower arrows denote gels on the 'C' strand (complementary to RNA). The tips of the arrows indicate the points to which the gels were read. Numbers in parentheses indicate the number of sequencing gels from each 5′-labelled starting point. When the distances that were read differed substantially between experiments (for example due to use of different secondary restriction sites) arrows with two tips are shown. The entire sequence was read on both strands except at the following points: restriction sites *Hin*fI 1,273, *Hpa*II 1,378 and *Alu*I 1,621 were sequenced only on S strand; sites *Xba*I 259 and *Eco*RI 1,596 were sequenced only on C strand. Secondary structure effects interfered with the reading of short segments on S strand at 244–7, 681–4, 1,146–7 and on C strand at 226–9, 666–7, 705–8. Methylated C residues in *Eco*RII sites were evident as gaps or very faint bands as follows: S4, C6; S804, C806; S1,180, C1,182. Gel 9 was unclear at 1,256 but at this and all other points referred to above, the sequence was unambiguously read on the other strand, and the restriction sites just mentioned were all experimentally demonstrated. (*Bst* NI cuts at methylated *Eco*RII sites.) In most cases the purification of fragments was straightforward, and sequencing was carried out after 5′ labelling and secondary restriction. At a few points more complex procedures were needed to allow the necessary overlapping sequences to be obtained. The relevant gels were as follows. S strand—gel 18, partial *Alu*I digest yielded fragment 1,094–1,262. 5′-terminal labelling was followed by secondary restriction at *Hin*fI 1,206. Gel 17, fragment bounded by *Hpa*II 1,469 and *Eco*RI 1,596, was labelled and subjected to partial *Sau*3A digestion to yield fragment terminating at 1,563. (Another *Sau*3A site at 1,497 is not shown.) Gel 8 (*Alu*I fragment 1,144–1,262), gel 9 (*Hin*fI fragment 1,206–1,273) and gel 16 (fragment bounded by *Hha*I 1,518 and *Eco*RI 1,596) were sequenced after strand separation. C strand—Gel 23, *Alu*I fragment 1,144–1,262, was labelled and subjected to secondary restriction at *Bst*NI 1,179; gel 22, *Hin*fI fragment 1,206–1,273, was subjected to secondary restriction with *Hae*III at 1,229 (not shown). Gels 13 (*Hin*fI fragment 1,273–1,348) and 19 (*Hpa*II fragment 1,378–1,469) were subjected to strand separation. (In this and some other experiments only one strand was successfully purified.) Gel 20, this crucial fragment (*Alu*I 1,262–1,621), which allowed sequencing through the *Eco*RI site, was obtained from the parent plasmid pX1r 101 by *Alu*I digestion of a 4.5-kilobase *Bam* HI fragment (these sites are shown in the top line of the figure). The *Alu*I fragment was labelled, subjected to secondary restriction at *Hpa*II 1,469 and the C strand was sequenced. The result confirmed the presence of a single *Eco*RI site and the correctness of the adjacent sequences as previously deduced from the subclones.

Figures 2 and 3 show a comparison of the *X. laevis* and *S. cerevisiae*[1] sequences. There are extensive tracts where the two sequences are almost identical. Most of the two sequences are >70% homologous, but there are also tracts with little or no homology. Almost all the *X. laevis* methylation sites are in regions of high homology with the yeast sequence. Four major regions of low homology are designated *a*, *b*, *c* and *d* in Fig. 3. Table 1 shows that these regions are considerably richer in C+G in *Xenopus* than in yeast.

Evolutionary trends

These findings provide insight into evolutionary trends in eukaryotic rRNA and suggest a general explanation for a well-known paradox: rRNAs from distantly related eukaryotes differ considerably in overall base composition[13,14] but show substantial homology by nucleic acid hybridization[15,16]. Clearly in *Xenopus* and yeast 18S rRNA the nucleic acid hybridization experiments show extensive, highly conserved sequences. The divergence in base composition between the two rRNAs results from cumulative changes in less highly conserved parts of the molecules, and much of this variability is concentrated in clearly circumscribed tracts. These range in size from short runs of a few nucleotides such as tract 65–75 (Fig. 2) to the longer, major variable regions of Table 1. Thus the pattern of sequence homology between *Xenopus* and yeast 18S rRNAs may be described as 'extensive but interrupted'.

How general is this pattern? Sequence analysis of the 3′ regions (rightwards from the *Eco*RI site) of 18S genes in silkworm[17] and fruitfly[18] shows that much of this region has also remained highly conserved in these organisms, but again the tract corresponding to region *d* is more variable. Recent refinements in nucleic acid hybridization experiments, in which heterologous rRNAs are hybridized to specific restriction fragments of rDNA, confirm the general pattern of 'extensive but interrupted' homology for both 18S and 28S rRNA[19,20].

How variable are the variable regions? Table 1 shows that the 18S variable regions in *Xenopus* and yeast differ from each other less widely than do the respective internal transcribed spacers. The regions differ in length only slightly between the two species, and although the *Xenopus* regions are rich in G+C, the corresponding yeast regions are, on average, no more rich in A+U than the rest of the sequence (*a* is rich in A+U but *b* is slightly rich in G+C even in yeast). This suggests that there are constraints on evolution in these 18S regions, but not such stringent ones as in the highly conserved regions. The 18S variable regions possibly play a semi-specific structural role in

Table 1 *X. laevis* and *S. cerevisiae* base composition data

		No. of nucleotides		U	A	C	G	%C+G
Complete 18S:								
X. laevis		1,825		411	432	467	515	53.8
S. cerevisiae		1,789		509	475	347	458	45.0
Major variable regions:								
X. laevis	*a*	103	(176–278)	16	20	38	29	65
S. cerevisiae		91	(177–267)	34	31	13	13	29
X. laevis	*b*	110	(651–760)	20	10	43	37	73
S. cerevisiae		106	(637–742)	35	12	29	30	56
X. laevis	*c*	72	(1,351–1,395, 1,411–1,437)	13	15	25	19	61
S. cerevisiae		71	(1,321–1,362, 1,378–1,406)	24	17	13	17	42
X. laevis	*d*	75	(1,695–1,769)	11	14	26	24	67
S. cerevisiae		71	(1,663–1,773)	17	19	15	20	49
X. laevis	Sum	360		60	59	132	109	67
S. cerevisiae		339		110	79	70	80	44
Rest of 18S:								
X. laevis		1,465		351	373	335	406	50.6
S. cerevisiae		1,450		399	396	277	378	45.2
ITS 1:								
X. laevis		557		19	69	230	239	84
S. cerevisiae		360		119	111	57	73	36

Numbers in parentheses indicate locations of the major variable regions in *X. laevis* and *S. cerevisiae*[1] 18S rRNA. Variable region *c* has two components separated by a short, conserved sequence with a methyl group in *Xenopus*. Data on ITS 1 are from refs 30 and 32.

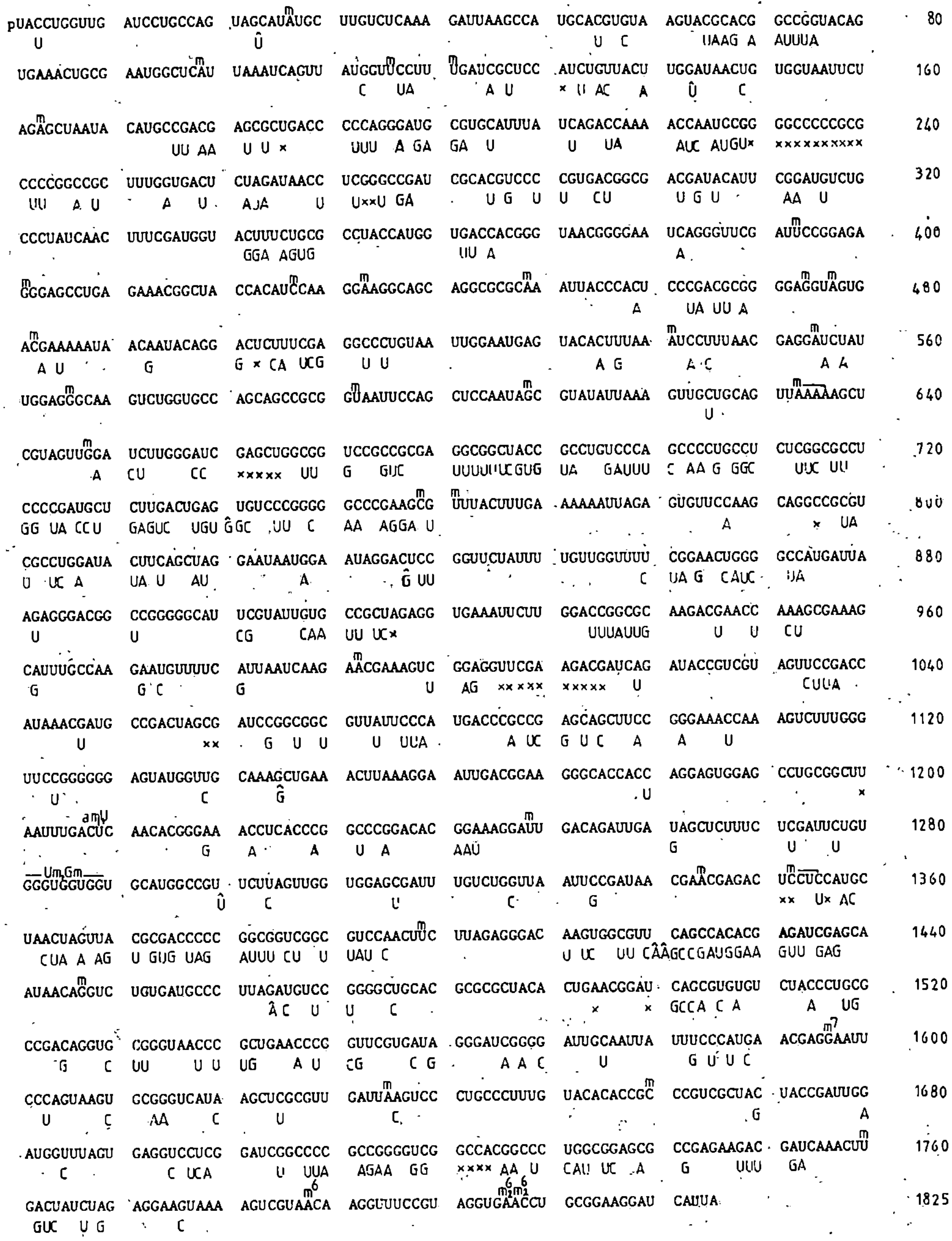

pUACCUGGUUG AUCCUGCCAG UAGCAUAUGC UUGUCUCAAA GAUUAAGCCA UGCACGUGUA AGUACGCACG GCCGGUACAG 80
UGAAACUGCG AAUGGCUCAU UAAAUCAGUU AUGGUUCCUU UGAUCGCUCC AUCUGUUACU UGGAUAACUG UGGUAAUUCU 160
AGAGCUAAUA CAUGCCGACG AGCGCUGACC CCCAGGGAUG CGUGCAUUUA UCAGACCAAA ACCAAUCCGG GGCCCCCGCG 240
CCCCGGCCGC UUUGGUGACU CUAGAUAACC UCGGGCCGAU CGCACGUCCC CGUGACGGCG ACGAUACAUU CGGAUGUCUG 320
CCCUAUCAAC UUUCGAUGGU ACUUUCUGCG CCUACCAUGG UGACCACGGG UAACGGGGAA UCAGGGUUCG AUUCCGGAGA 400
GGGAGCCUGA GAAACGGCUA CCACAUCCAA GGAAGGCAGC AGGCGCGCAA AUUACCCACU CCCGACGCGG GGAGGUAGUG 480
ACGAAAAAUA ACAAUACAGG ACUCUUUCGA GGCCCUGUAA UUGGAAUGAG UACACUUUAA AUCCUUUAAC GAGGAUCUAU 560
UGGAGGGCAA GUCUGGUGCC AGCAGCCGCG GUAAUUCCAG CUCCAAUAGC GUAUAUUAAA GUUGCUGCAG UUAAAAAGCU 640
CGUAGUUGGA UCUUGGGAUC GAGCUGGCGG UCCGCCGCGA GGCGGCUACC GCCUGUCCCA GCCCCUGCCU CUCGGCGCCU 720
CCCCGAUGCU CUUGACUGAG UGUCCCGGGG GCCCGAAGCG UUUACUUUGA AAAAAUUAGA GUGUUCCAAG CAGGCCCGCGU 800
CGCCUGGAUA CUUCAGCUAG GAAUAAUGGA AUAGGACUCC GGUUCUAUUU UGUUGGUUUU CGGAACUGGG GCCAUGAUUA 880
AGAGGGACGG CCGGGGGCAU UCGUAUUGUG CCGCUAGAGG UGAAAUUCUU GGACCGGCGC AAGACGAACC AAAGCGAAAG 960
CAUUUGCCAA GAAUGUUUUC AUUAAUCAAG AACGAAAGUC GGAGGUUCGA AGACGAUCAG AUACCGUCGU AGUUCCGACC 1040
AUAAACGAUG CCGACUAGCG AUCCGGCGGC GUUAUUCCCA UGACCCGCCG AGCAGCUUCC GGGAAACCAA AGUCUUUGGG 1120
UUCCGGGGGG AGUAUGGUUG CAAAGCUGAA ACUUAAAGGA AUUGACGGAA GGGCACCACC AGGAGUGGAG CCUGCGGCUU 1200
AAUUUGACUC AACACGGGAA ACCUCACCCG GCCCGGACAC GGAAAGGAUU GACAGAUUGA UAGCUCUUUC UCGAUUCUGU 1280
GGGUGGUGGU GCAUGGCCGU UCUUAGUUGG UGGAGCGAUU UGUCUGGUUA AUUCCGAUAA CGAACGAGAC UCCUCCAUGC 1360
UAACUAGUUA CGCGACCCCC GGCGGUCGGC GUCCAACUUC UUAGAGGGAC AAGUGGCGUU CAGCCACACG AGAUCGAGCA 1440
AUAACAGGUC UGUGAUGCCC UUAGAUGUCC GGGGCUGCAC GCGCGCUACA CUGAACGGAU CAGCGUGUGU CUACCCUGCG 1520
CCGACAGGUG CGGGUAACCC GCUGAACCCC GUUCGUGAUA GGGAUCGGGG AUUGCAAUUA UUUCCCAUGA ACGAGGAAUU 1600
CCCAGUAAGU GCGGGUCAUA AGCUCGCGUU GAUUAAGUCC CUGCCCUUUG UACACACCGC CCGUCGCUAC UACCGAUUGG 1680
AUGGUUUAGU GAGGUCCUCG GAUCGGCCCC GCCGGGGUCG GCCACGGCCC UGGCGGAGCG CCGAGAAGAC GAUCAAACUU 1760
GACUAUCUAG AGGAAGUAAA AGUCGUAACA AGGUUUCCGU AGGUGAACCU GCGGAAGGAU CAUUA 1825

Fig. 2 Nucleotide sequence of *X. laevis* 18S rRNA, with positions of methyl groups indicated by superscripts and comparison with the yeast *S. cerevisiae*[1] sequence indicated by subscripts. Most of the methyl groups were located from T_1 oligonucleotide data[6–9,11], but in several cases additional data from pancreatic ribonuclease fingerprints (B. E. H. M., unpublished results) were also needed—this work will be fully reported elsewhere. The methylation sites are as follows: m superscripts denote 2′-*O*-methyl groups at 27, 99, 116, 121, 163 (not 181 as was reported previously[2]), 393, 401, 427, 433, 449, 474, 477, 482, 541, 555, 566, 592, 609, one of 633–636, 648, 759, 761, 992, 1,249, Um and Gm between 1,280–1,289, 1,344, 1,352 or 1,355, 1,399, 1,447, 1,635, 1,660 and 1,760. Remaining superscripts denote base methyl groups at amψ 1,209, m^7G 1,596 (encoded at first base of *Eco*RI site), m^6A 1,788, m^6_2A 1,806 and 1,807. This list includes all known *X. laevis* 18S methyl groups except for one fractional 2′-*O*-methyl U which occurs at an unidentified location in the 5′ third of the molecule. Also mentioned in the text are oligonucleotides containing oligo(A)tracts[3]. These are: A_4C, 218–222; A_5U, 484–489; $(A_3, AmA)G$, 633–638; $A_6\psi$, 770–776; and A_4G, 1,778–1,782. Differences between the *Xenopus* and yeast sequences are indicated by plain subscripts in the case of substitutions; additional nucleotides in the yeast sequence are shown as subscripts between nucleotides in the *Xenopus* sequence and denoted by a circumflex sign. Crosses indicate deletions. In regions of low homology deletions are arbitrarily placed to allow optimal alignment between regions of good homology. (One apparent deletion from the yeast sequence occurs in a region of otherwise good homology with *Xenopus*, in nucleotides 1,006–1,015 of the latter. This part of the yeast sequence determination used a *Sau*3A site which does not seem to have been overlapped during analysis[1].)

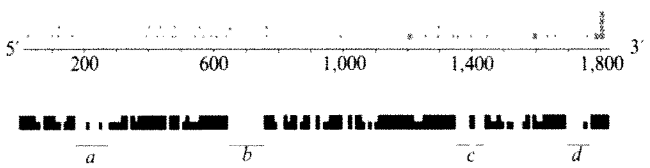


Fig. 3 Location of methylation sites in *X. laevis* 13S rRNA and pattern of homology with *S. cerevisiae*. Upper section: crosses denote 2′-*O*-ribose methyl groups; circled crosses denote base methyl groups. Lower section: high blocks indicate regions of *X. laevis* sequence possessing 85–100% homology with *S. cerevisiae*; low blocks indicate 70–85% homology. *a*, *b*, *c* and *d* denote major variable regions (see text and Table 1).

the ribosome. In 28S rRNA two long, variable regions do differ substantially in length between species[21,22].

What is the significance of localization of methyl groups within the conserved regions? Clearly many sites within these conserved regions provide recognition signals for methylases during ribosome maturation. However, further interpretation is difficult because the conservation of neither sequence nor methyl group numbers is absolute. (Yeast 18S rRNA possesses the same base methyl groups as *Xenopus* but some 15 fewer 2′-*O*-methyl groups[23,24].) Determination of the methylation sites in the yeast sequence should help to clarify the relationship between methylation and primary structure.

In contrast to the readily apparent and extensive tracts of homology between the *X. laevis* and yeast 18S sequences, we can detect little major homology with the *Escherichia coli* (prokaryotic) 16S sequence[25,26] except near the 3′ end. This lack of extensive sequence homology between eukaryotic 18S and prokaryotic 16S rRNA is in agreement with results obtained from nucleic acid hybridization[15,19] and T_1 ribonuclease oligonucleotide data[27]. The relatively high degree of homology between the *Xenopus* and yeast sequences together with the lack of comparable homology between either of these and *E. coli* 16S rRNA further supports the widely held view that the divergence between the eukaryotic and prokaryotic systems for gene expression was a very ancient one[27–29].

Conclusions

Our conclusions and inferences on the structure and evolution of eukaryotic rRNA are as follows. (1) Substantial parts of the sequences have remained highly conserved since the last common ancestor of metazoa and yeasts, a period of probably 10^9 yr. (2) Crucial events in the function or assembly of eukaryotic ribosomes, including rRNA methylation, centre on these conserved regions. (3) Tracts where primary structure is phylogenetically more variable are interspersed among the conserved regions. (4) The most clearly defined variable tracts are rich in C+G in *Xenopus*, and probably in vertebrates generally. (5) Phylogenetic change in these variable regions is less extreme than in the transcribed spacer regions.

We thank R. Reeder for a gift of the clone pX1r 101, Lucinda Hall and W. Darlington for discussion, J. Forbes and Mary Robertson for technical assistance and the MRC for financial support. M.S. was on study leave from the Department of Biology, Quaid-I-Azam University, Islamabad, Pakistan.

Received 5 January; accepted 18 March 1981.

1. Rubstov, P. M. *et al. Nucleic Acids Res.* **8,** 5779–5794 (1980).
2. Salim, M. & Maden, B. E. H. *Nucleic Acids Res.* **8,** 2871–2884 (1980).
3. Khan, M. S. N. & Maden, B. E. H. *J. molec. Biol.* **101,** 235–254 (1976).
4. Boseley, P. G., Tuyns, A. & Birnstiel, M. L. *Nucleic Acids Res.* **5,** 1121–1137 (1978).
5. Boseley, P. G., Moss, T., Machler, M., Portmann, R. & Birnstiel, M. L. *Cell* **17,** 19–31 (1979).
6. Khan, M. S. N., Salim, M. & Maden, B. E. H. *Biochem. J.* **169,** 531–542 (1978).
7. Maden, B. E. H. & Salim, M. *J. molec. Biol.* **88,** 133–164 (1974).
8. Fuke, M. & Busch, H. *Nucleic Acids Res.* **7,** 1131–1135 (1979).
9. Choi, Y. C. & Busch, H. *Biochemistry* **17,** 2551–2560 (1978).
10. Maden, B. E. H. (in preparation).
11. Maden, B. E. H. *Nature* **288,** 293–296 (1980).
12. Saponara, A. G. & Enger, M. D. *Biochim. biophys. Acta* **349,** 61–77 (1974).
13. Amaldi, F. *Nature* **221,** 95–96 (1969).
14. Attardi, G. & Amaldi, F. *A. Rev. Biochem.* **39,** 183–226 (1970).
15. Sinclair, J. & Brown, D. D. *Biochemistry* **10,** 2761–2769 (1971).
16. Gerbi, S. A. *J. molec. Biol.* **106,** 791–816 (1976).
17. Samols, D. R., Hagenbuchle, O. & Gage, L. P. *Nucleic Acids Res.* **7,** 1109–1119 (1979).
18. Jordan, B. R., Latil-Damotte, M. & Jourdan, R. *FEBS Lett.* **117,** 227–231 (1980).
19. Gourse, R. L. & Gerbi, S. A. *J. molec. Biol.* **140,** 321–339 (1980).
20. Cox, R. A. & Thompson, R. D. *Biochem. J.* **187,** 75–90 (1980).
21. Wellauer, P. K., Dawid, I. B., Kelley, D. E. & Perry, R. P. *J. molec. Biol.* **89,** 397–407 (1974).
22. Schibler, U., Wyler, T. & Hagenbuchle, O. *J. molec. Biol.* **94,** 503–517 (1975).
23. Brand, R. C., Klootwijk, J., Van Steenbergen, T. J. M., De Kok, A. J. & Planta, R. J. *Eur. J. Biochem.* **75,** 311–318 (1977).
24. Klootwijk, J. & Planta, R. J. *Eur. J. Biochem.* **39,** 325–333 (1973).
25. Brosius, J., Palmer, M. L., Poindexter, J. K. & Noller, H. F. *Proc. natn. Acad. Sci. U.S.A.* **75,** 4801–4805 (1978).
26. Carbon, P., Ehresmann, C., Ehresmann, B. & Ebel, J. P. *FEBS Lett.* **94,** 152–156 (1978).
27. Woese, C. R. & Fox, G. E. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5088–5090 (1977).
28. Doolitte, R. F. *Nature* **272,** 581–582 (1978).
29. Darnell, J. E. *Science* **202,** 1257–1260 (1978).
30. Hall, L. M. C. & Maden, B. E. H. *Nucleic Acids Res.* **8,** 5993–6005 (1980).
31. Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65,** 499–560 (1980).
32. Skryabin, K. G., Krayev, A. S., Rubstov, P. M. & Bayev, A. A. *Dokl. Akad. Nauk SSSR* **247,** 761–765 (1979).

LETTERS

RG0044−2958: a peculiar M supergiant at a distance of 2.5 Mpc

Gerard Gilmore

Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK

I. Neill Reid

Department of Astronomy, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ, UK

Peculiar M supergiants similar to VX Sgr and VY CMa are characterized by very red visual colours, a flux excess near 9,000 Å (*I* band) and a substantial excess longwards of 2μm (refs 1, 2). HV11417 has recently been reported to be the first similar star discovered in the Small Magellanic Cloud[3]. We report here the discovery of another star of this type, RG0044−2958, which is more unusual in that it is located at a distance of ~2.5 Mpc, and is not associated with a detectable galaxy. It is probably the brightest star in a very low surface brightness dwarf galaxy in the Sculptor group.

We are presently studying the stellar distribution towards the South Galactic Pole to $m \approx 20$ in the B, V, R and I passbands, using UK Schmidt Telescope plates measured on the COSMOS facility at the Royal Observatory, Edinburgh. Our primary aim is the determination of the intrinsically faint end of the stellar luminosity function, and the solar neighbourhood space density of low mass stars. We use as a sensitive absolute magnitude calibrator for main sequence dwarfs the $M_V/V-I$ relationship derived directly from stars of accurately known parallax[4]. The resulting luminosity function can then be compared directly with that derived for the same field using classical kinematic techniques[5]. We are also observing the very red stars in the near IR (*JHK*) to allow luminosity class (dwarf/giant) and population type (halo/disk) discrimination.

One of the more interesting objects we have studied has the broadband energy distribution shown in Fig. 1, which also gives the energy distribution for the similar object HV11417 (from ref. 3), the two very late M dwarfs, VB8 and RG0050−2722 (ref. 6), and a 'typical' giant and supergiant[7]. The ability of *JHK* photometry to provide dwarf–giant classification is evident, while the substantial '*I*' band excess shows the object to be a peculiar M supergiant similar to HV11417 and VX Sgr. This identification has been confirmed from a spectrum obtained with

0028–0836/81/210208—03$01.00

Table 1 Magnitudes for RG0044 − 2958

Passband	Mag
B	22.4 ± 0.4
V	20.5 ± 0.2
R	19.1 ± 0.2
I	17.0 ± 0.2
J	16.83 ± 0.04
H	15.86 ± 0.03
K	15.47 ± 0.04

the multichannel scanner on the Palomar 5-m telescope (J.B. Oke and R. D. Cannon, personal communication).

There are two major differences between RG0044 − 2958 and VX Sgr: RG0044-2958 is not associated with a detectable galaxy, at least to the limit of the SRC '*J*' Sky Survey, which has a limiting magnitude near $m_J = 23$, and it has not shown any substantial variability since 1976.7. There is, however, a strong 1-yr selection effect in our data, and any variability with a period near 1 yr would not have been detected. Useful images of RG0044 − 2958 appear on 15 USKT plates, of which five are IIIa-J(B), three are IIIa-F(R), four are V band and three are I band. It is not visible on either POSS print or the ESO 'B' Survey plate. The derived magnitudes are shown in Table 1, together with our IR (*JHK*) data. These were obtained with the 3.9-m Anglo-Australian Telescope on 27 October 1980. One of our deep I plates was obtained the following night, and precludes extreme variability as an explanation of the I/R colours. If RG0044 − 2958 does prove to be non-variable, the suggestion[3] that stars of this type are extreme long-period variables rather than peculiar supergiants would be disproved. The variable nature of the first three known stars of this type is readily explained as a selection effect, as the visual colours are fairly normal (for such late-type stars) and would not excite spectroscopic interest, whereas variability might. As seen in Fig. 1, either visual spectroscopy or both I band and near IR photometry are necessary to discover the peculiar nature of these stars.

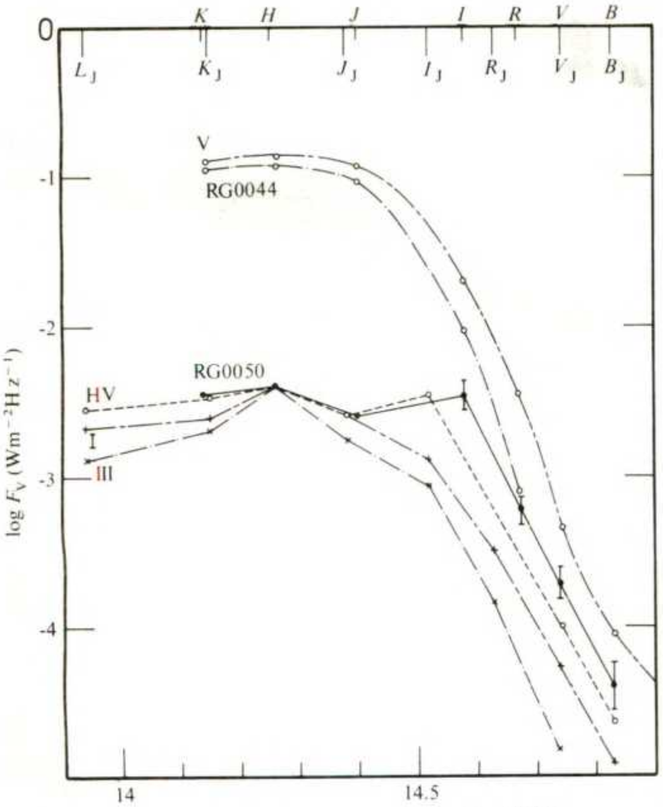


Fig. 1 Spectral energy distributions for the peculiar M supergiants RG0044 − 2958 (—), HV11417 (ref. 3) (HV, – – –), and two late-type M dwarfs, VB8 (ref. 6) (V, - - - -) and RG0050 − 2722 (ref. 6) (— · — ·). A typical M5 giant (III) and supergiant (I) from the data of Lee[7] are also shown. (The zero points are arbitrary.) The effective wavelengths of our filter passbands (*BVRIJHK*) are marked at the top of the figure and are appropriate for our observations (RG0044, RG0050 and VB8). Those of the Johnson photometric system are marked with a subscript J (*BVRIJKL*) and apply to the observations of curves HV, I and III.

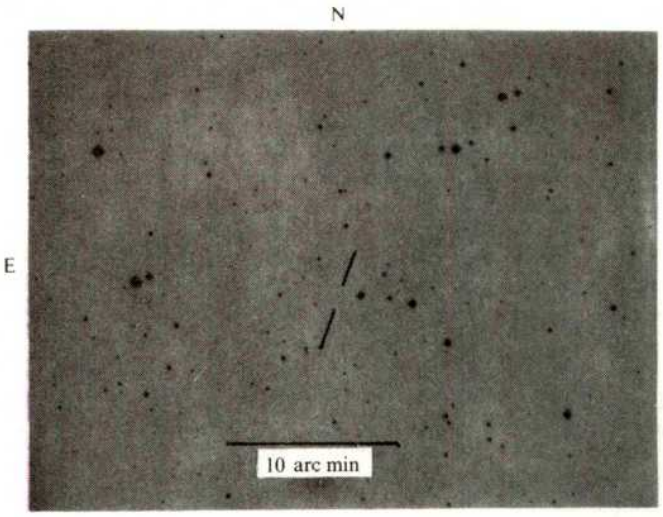


Fig. 2 A finding chart for RG0044 − 2958 reproduced from an I band UKST plate. North-east is at the top left corner, and the scale is marked. (Reproduced with permission from the UK Schmidt Telescope Unit, Royal Observatory, Edinburgh.)

The absolute magnitude of RG0044 − 2958 cannot be directly determined from our available data. Elias *et al.*[3] derive a visual absolute magnitude $M_V = -5.1$ for HV11417, and note that the corresponding value for VX Sgr is only poorly known. Humphreys (Fig. 3 of ref. 8) shows the late M supergiants in M33 to have typically $M_V = -6.8$, with some as bright as $M_V = -8$. If RG0044 − 2958 is similar to these stars, it must have a distance modulus of some 27 mag, although this value may be uncertain by several magnitudes. Significantly, the Sculptor group of galaxies lies in the same direction in the sky, and has distance modulus 27.2 ± 0.2 mag with little if any reddening[9]. If this star is, in fact, associated with this group, its absolute magnitude becomes $M_V = -6.6 \pm 0.5$, consistent with the expected value. However, no associated galaxy is visible. The nearest Sculptor group galaxy is NGC253 (ref. 10), which is some 4.25° away. At a distance of 2.5 Mpc, this corresponds to a spatial separation of ~200 kpc. A velocity of $\sim 0.02c$ is then required for this star to have escaped from NGC253 in the lifetime of the precursor of an M supergiant. It is more likely that RG0044 − 2958 is the brightest star in a dwarf galaxy of low surface brightness, which is associated with the Sculptor group and possibly with NGC253 itself.

The position of RG0044 − 2958, determined relative to a grid of 76 SAO stars from the COSMOS measurements, is RA00 h 44 min 26.2 s, dec −29°58′49″ (1950.0) and a finding chart is provided in Fig. 2.

As Fig. 1 shows there is no detectable difference between an extreme M supergiant and a very late M dwarf from broadband visual colours. The true number of objects similar to RG0044 − 2958 therefore cannot be determined until we have completed *JHK* photometry of our sample of very red stars.

We thank J. B. Oke and R. D. Cannon for obtaining the spectrum with the 5-m Hale Telescope on Palomar Mountain, and the UKST and COSMOS groups at the Royal Observatory, Edinburgh, for assistance.

Received 4 February; accepted 23 March 1981.

1. Humphreys, R. M. *Astrophys. J.* **188,** 75 (1974).
2. Fawley, W. M. *Astrophys. J.* **218,** 181 (1977).

3. Elias, J. H., Frogel, J. A. & Humphreys, R. M. *Astrophys. J. Lett.* **242**, L13 (1980).
4. Strand, K. Aa *et al. IAU Symp. No.* 61 159 (1974).
5. Murray, C. A. & Corben, P. M. *Mon. Not. R. astr. Soc.* **187**, 723 (1979).
6. Reid, I. N. & Gilmore, G. F. *Mon. Not. R. astr. Soc.* (in the press).
7. Lee, T. A. *Astrophys. J.* **162**, 217 (1970).
8. Humphreys, R. M. *Astrophys. J.* **241**, 587 (1980).
9. de Vaucouleurs, G. *Astrophys. J.* **224**, 710 (1978).
10. Pence, W. D. *Astrophys. J.* **239**, 54 (1980).

Solar radio blips and X-ray kernels

A. O. Benz*, E. Fürst†, W. Hirth‡
& M. R. Perrenoud*

* Institute of Astronomy, ETH, Hochstrasse 58, CH-8044 Zurich, Switzerland
† Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, D-5300 Bonn, FRG
‡ Radioastronomisches Institut der Universität Bonn, Auf dem Hügel 71, D-5300 Bonn, FRG

Hudson[1] proposed that the primary flare energy release goes into fast electrons. Others[2,3] found the flare soft X-ray emission to consist of bright, small kernels (knots and loops) embedded in a more diffuse halo. These kernels with a typical size of 5′ arc s and a temperature of 6–12 × 10⁶K emit in radio waves at 3.7 and 11 cm (ref. 4). The discovery of a radio brightness temperature equal to the temperature derived from the X rays and the radio size larger than in X rays shows that the observed radio emission is thermal (free–free) and the source optically thick at λ > 3.7 cm. Therefore, Kundu *et al.*[4] did not observe the fast electrons expected to heat the kernels. Here we present high time resolution measurements at 2.8 cm from the Effelsberg 100-m telescope and at decimetric and metric wavelengths with the ETH Zurich spectrometers. Both instruments have detected signatures of non-thermal particles.

We observed the microwave flux of McMath region 15856 on 2 March 1979. The Effelsberg telescope at 10.69 GHz yields a half-power beam width of 1.25 arc min. Simultaneous observations in the metric band by the Nançay heliograph and the Zurich digital spectrometer did not receive any signal in the 2-h observing period. However, decimetric registrations with the Zurich analogue spectrograph showed four groups of short, narrow-band spikes which coincided with the four largest microwave peaks of the observing period. This activity was also associated with four weak soft X-ray bursts as observed with the GOES 2 satellite. Figure 1 summarizes the data for one example. Note the following prominent properties of the radio emissions: (1) The time constant in decimetric and high-frequency microwaves is very small (2–3 s). (2) The radio bursts occur mainly in the rise phase of the soft X-ray emission (delay time of X-ray maximum 3–30 s). (3) The correlation between microwaves and decimetric radiation is excellent. (4) The amplitude of the microwave bursts is of the order of 2 SFU, much smaller than the bursts commonly observed by small telescopes monitoring the full solar disk and routinely reported in the flare bulletins.

These properties apply to all observed events and seem to be typical. All these events look very similar in microwave and X rays, but their shapes in the decimetric frequency–time plane show considerable variation. The flux ratio between single associated peaks in decimetre and microwaves fluctuates in a large range. The microwave emission has unusually short time scales and a very small amplitude. The bandwidth in decimetre waves is much smaller than commonly observed for type III bursts. These differences have led us to coin the the name 'blip' for the radio emission.

No other activity has occurred in the decimetric and soft X-ray bands during that period. A chance coincidence is out of question. In microwaves, however, many smaller peaks have been observed with similar characteristics. Of the four blips to be described here, three coincided with the maxima of subflares in McMath 15856.

A temperature of 10×10^6 K is derived for the plasma of the soft X-ray source from the ratio of the 1–8 Å and 0.5–4 Å intensities (correcting for line emission). The emission measure is 1.2×10^{48} cm^{-3}. X-ray kernels have the tendency to avoid flares with non-thermal radio bursts[2,3]. However, because the observed radio emission does not resemble impulsive radio bursts of conventional types and fluxes, we consider it very likely that the X-ray emission originates in kernels. This is supported by the fact that the observed temperature, emission measure and the correlation with subflares is exactly as in soft X-ray kernels[2]. Thus we adopt a source size of 5 × 10 arc s and a volume of 10^{26} cm^3. This yields a density of $\sim 10^{11}$cm^{-3}.

The optical depth of the kernel at 10.7 GHz is 0.19, and the thermal radiation amounts to 0.67 SFU. Because the observed peaks are a factor of 2 or 3 higher and do not correlate with the X-ray burst, we conclude that the thermal radiation is absorbed in the vicinity of the source. A non-thermal origin for the microwave blips is also suggested from their correlation with their decimetric counterpart which, by its spectrum, is certainly not thermal. Thus we consider non-thermal electrons with velocities β (in units of the speed of light) estimated from the collisional loss time t_c, where

$$t_c = 3.3 \times 10^{12} \frac{(1-\sqrt{1-\beta^2})^2}{n\beta\,(1-\beta^2)} \approx 3\text{s}$$

with the ion density n in cm^{-3}. From the observed decay time of ~3 s one derives $\beta = 0.4$ corresponding to 50-keV electrons. The observed energy input of 2.76×10^{28} erg observed in X rays requires a total number of 4×10^{35} fast electrons. If the X-ray source is heated by the collisional energy loss of these electrons, a total number, N_T, of 4×10^{35} is required.

To estimate the gyrosynchrotron radiation of these electrons in a magnetic field B we applied the emission formula given by

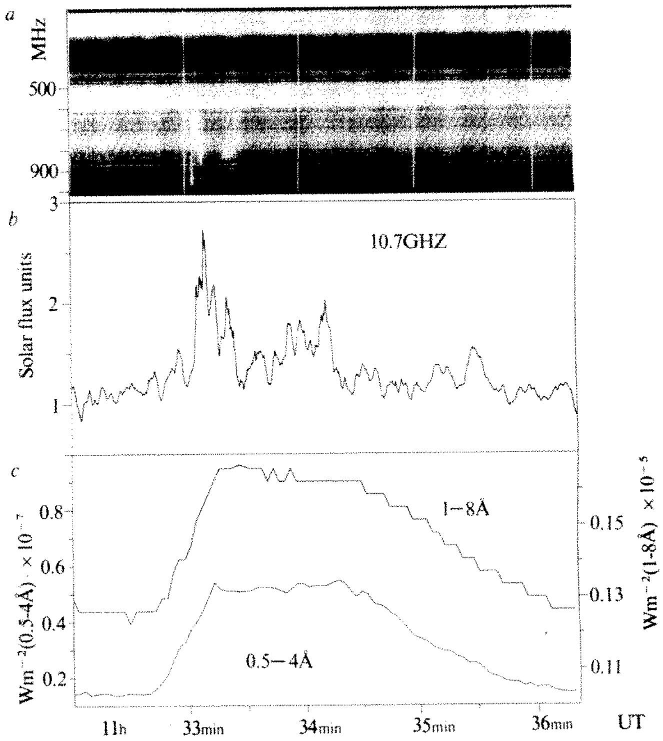


Fig. 1 Solar blip on 2 March 1979 in different wave bands. *a*, Spectrogram of the Zurich analogue spectrometer showing flux (bright) versus frequency (down) and time (right). Horizontal straight lines are all instrumental. The vertical lines are minute marks. The threshold is ~30 SFU. *b*, Total flux versus time observed by the 100-m telescope in Effelsberg (MPIfR, Bonn) of McMath region 15856. *c*, GOES 2 soft X-ray measurements in the bands 0.5–4 Å (left scale) and 1–8 Å (right scale).

0028-0836/81/210210—02$01.00

Bekefi[5] (formula 6.14). The number N of fast electrons present at any one time is their total number, N_T, times the ratio of the decay time and the total blip activity duration ($N=6\times10^{34}$). From the observed average flux of about 0.5 SFU and from the estimated free–free absorption (source $\tau=0.19$; foreground $\tau=1.5$, as required from the lack of thermal radiation) we arrive at $B=250$–300 G for low pitch angles and large angles (>60°) between the line of sight and magnetic field direction.

In the total number of electrons estimated above we have not included fast particles that avoid energy loss in the X-ray source by escape to regions of lower density. Only a small fraction, however, is needed to produce the observed decimetric blip presumably by some plasma emission process.

In conclusion, according to the above discussion the microwave signal of non-thermal electrons is not to be expected at frequencies below 10 GHz. Therefore the previous observations by Kundu *et al.*[4] probably refer to thermal emission of the hot plasma and not to fast particles. The discovery of decimetric signatures proves the presence of non-thermal particles in at least the larger microwave blips. Two important questions remain: first, is the observed rate of one blip (above 1 SFU at 10.7 GHz) in 25 min typical? Second, are the smaller fluctuations[6,7], which occur all the time, of the same nature? An extensive study involving major radio telescopes and the SMM satellite is in progress.

We acknowledge observational support by the Nançay group (M. Pick) and the financial support of the Swiss NSF for the construction of the Zurich radio spectrometers. E. F. and W. H. were supported by the Deutsche Forschungsgemeinschaft.

Received 30 January; accepted 23 March 1981.

1. Hudson, H. S. in *High Energy Phenomena on the Sun* (eds Ranecty, R. & Stone, R. G.) 207 (NASA SP-342, 1973).
2. Kahler, S. W., Petrasso, R. D. & Kane, S. R. *Solar Phys.* **50,** 179 (1976).
3. Vorpahl, G. A., Gibson, E. G., Landecker, P. B., McKenzie, D. L. & Underwood, G. H. *Solar Phys.* **45,** 199 1975).
4. Kundu, M. R., Alissandrakis, C. E. & Kahler, S. W. *Solar Phys.* **50,** 429 (1976).
5. Bekefi, G. *Radiation Processes in Plasmas,* 180 (Wiley, New York, 1966).
6. Butz, M., Hirth, W. & Fürst, E. *Astr. Astrophys.* **72,** 211 (1979).
7. Benz, A. O. & Fitze, H. R. *Astr. Astrophys.* **76,** 354 (1979).

New positive ion species in the stratosphere

G. Henschen & F. Arnold
Max-Planck-Institut für Kernphysik, 6900 Heidelberg, FRG

Our understanding of atmospheric ions in the ionospheric region is far better than of those in the denser regions, the stratosphere and the troposphere. Mass spectrometric measurements conducted on rockets[1] and balloons[2–4] at altitudes above ~35 km recently revealed that the major ions are proton hydrates $(H^+(H_2O)_n$ and heteromolecular cluster ions identified as $H^+X_L(H_2O)_M$ (refs 2, 5) (X-ions), which probably are formed from proton hydrates by reactions involving a stratospheric trace gas X (ref. 2). The most likely candidate for X seems to be acetonitrile (CH_3CN) (refs 2, 6). The abundance ratio of X ions and proton hydrates increases with decreasing altitude from ~0.02 around 50 km to ~2 around 35 km. It has only been possible to detect ion species with number densities larger than about 6 cm^{-3}, corresponding to a fractional ion abundance of 3% around 35 km. We report here new balloon-borne positive-ion composition measurements around 34 km using improved mass spectrometers which allowed for an ion detection limit of ~0.02 cm^{-3} or 0.1%.

The instruments used have been discussed in detail elsewhere[2,5]. Improving the ion injection into the quadrupole mass filter and the efficiency of ion detection by the channeltron has increased the sensitivity. Thus, the overall ion transmission of the instrument was increased to ~20% depending on the mass resolution setting. Usually, an increase of mass resolution causes a decrease of ion transmission, particularly at higher masses. Decreasing the dark count rate of the channeltron from ~0.02 to 0.001 s^{-1} contributed to an increase of the signal-to-noise ratio and thereby to the sensitivity for ion detection, as the channeltron is operated in a saturated pulse counting mode. For a sufficiently large integration time (~10^5 s) a minimum detection limit of ~0.002 ions cm^{-3} or 0.0001% can be reached for average mass resolution. Considering a more realistic integration time of only 10^3 s the detection limit becomes ~0.2 ions cm^{-3} or 0.01%. If a more realistic minimum number of ion counts per mass peak of 10 rather than 1 were considered, the detection limit becomes ~2 ions cm^{-3} corresponding to a fractional ion abundance of 0.1%. When compared with detection limits reached in stratospheric ion composition measurements reported previously, this value is lower by a factor of ~30.

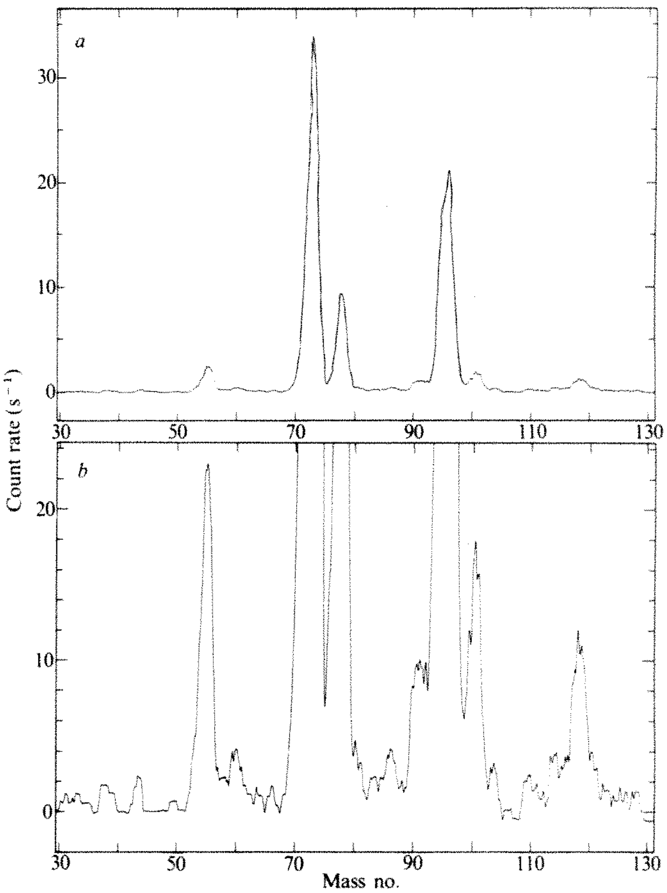


Fig. 1 *a*, Positive ion mass spectrum obtained at an altitude of 34 km (flight B 14). *b*, The same spectrum at an expanded scale for relative count rates. A relatively low mass resolution was used to reach higher sensitivity.

Three balloon-borne positive ion composition measurements were made around 34 km. Characteristic data of the experiments are given in Table 1. The mass range of the instruments which extended from 0 to 280 AMU was scanned in ~11 s to allow for a high time resolution. Data measured at the same altitude were integrated during the data reduction by superimposition of a large number of individual mass scans.

The observed mass peaks may be divided into three groups having fractional count rates in the ranges: (1) 10–100%; (2) 1–10%; and (3) 0.1–1%. Group (1) includes the peaks 73, 78 and 96 AMU for all three flights; of these, either 73 or 78 dominate. Group (2) includes 55, 60, 91, 101 and 119 AMU. Among these, peaks 60 and 91 do not exceed the 1% level in all flights. All ions belonging to group (1) and (2) have already been detected previously and were identified as members of the ion families $H^+(H_2O)_n$ (55, 73, 91), $H^+X(H_2O)_l$ (60, 78, 96), and $H^+X_2(H_2O)_l$ (101, 119) (refs 1, 2, 5).

Table 1 Measured minor ion species having fractional count rates <1% and not belonging to the major ion families $H^+(H_2O)_n$ and $H^+X(H_2O)_n$

Ion	B9	B10	B14	$A^+(H_2O)$	$A^+(H_2O)_2$	$A^+(H_2O)_3$	$A^+X(H_2O)$	$A^+X(H_2O)_2$	A^+
45	+	+		$\underline{27}$	9	—	—	—	27
49	+	+		$\underline{31}$	13	—	—	—	31
58			+	40	$\underline{22}$	4	—	—	22
63	+	+	+	45	$\underline{27}$	9			27
67	+	+	+	49	$\underline{31}$	13	8	—	31
81	+	+	+	63	45	$\underline{27}$	22	4	27
86	+	+	+	68	50	$\underline{32}$	27	9	32
104	+	+	+	86	$\underline{68}$	50	45	$\underline{27}$	27
110	+	+	+	92	$\underline{74}$	56	51	$\underline{33}$	33
117			+	99	81	$\underline{63}$	58	40	63
122	+		+	104	86	$\underline{68}$	$\underline{63}$	$\underline{45}$	63
125			+	107	89	71	68	$\underline{48}$	48
128		+		110	92	$\underline{74}$	69	51	74

Mass numbers for ion cores A^+ are also given assuming that the ions have the form $A^+(H_2O)_n$ or $A^+X(H_2O)_n$ with X having mass 41. The most probable core ions A^+ are underlined. All mass numbers given have an uncertainty of ±1 AMU. The balloon flights B9, B10 and B14 took place over southern France on 30 April, 21 June and 26 September 1980 and reached float altitudes of 33, 34 and 34 km.

Several smaller peaks falling into group (3) can also easily be identified. They include 37 ± 1, 42 ± 1, 45 ± 1, 49 ± 1, 58 ± 1, 83 ± 1, 86 ± 1, 104 ± 1, 110 ± 1 and 114 ± 1. The mass identification for these peaks is less certain as their shapes are not always well defined due to a small number of total ion counts contained in the peak. In other cases, small peaks are located besides large peaks and therefore appear only as shoulders. An example for a mass spectrum (flight B 14) it shown in Fig. 1. A mass uncertainty of ±1 AMU must therefore be allowed for. Besides these 10 ions more members of group (3) were apparently detected in particular in the mass regions 60–73, 78–91 and 110–130. These peaks are only poorly defined and therefore uncertain. The count rates in these regions, however, are significantly above the background signal. Table 1 lists some peaks belonging to group (3) (out of 16) and indicates the flights in which the peaks were observed. Figure 2 shows the measured fractional ion count rates for the major families: (1) $H^+(H_2O)_n$; (2) $H^+X(H_2O)_n$; and (3) $H^+X_2(H_2O)_n$. The average numbers of ligand molecules are 2.7, 2.1, and 1.9. Evidently, incorporating X into the cluster ions decreases the average ligand number a result consistent with previous stratospheric measurements[5] and laboratory studies[6,7].

The total fractional count rates for the above ion families are 48.2, 42.0 and 5.9%. Considering that families (2) and (3) are formed by incorporation of X into ions of families (1) and (2), the incorporation of a second X seems much less efficient than that of the first X. This contrasts with our previous measurement[2,5] which revealed that the efficiency for incorporation of one and two X molecules is about the same while that for the third X is much lower. The latter was interpreted in terms of backswitching of $H^+X_3(H_2O)_n$ ions with H_2O or of thermal detachment of an X from these ions[5,6]. The present finding may suggest that due to the higher stratospheric temperature (240 K) during the B14 flight these processes even affected $H^+X_2(H_2O)_n$ ions.

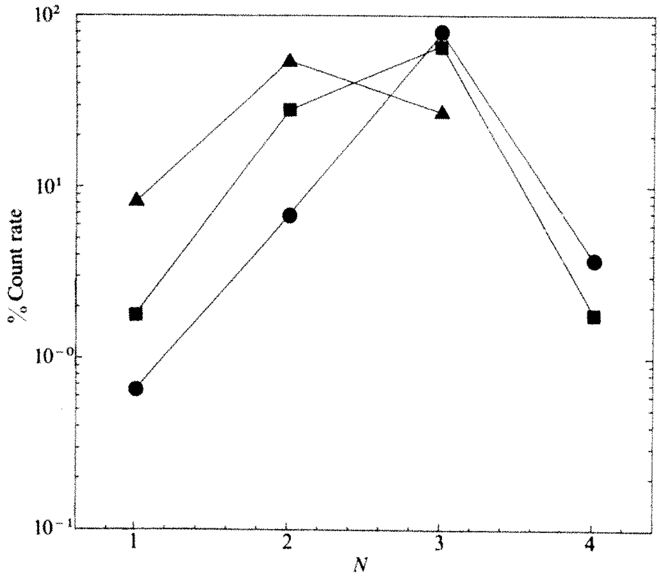


Fig. 2 Per cent count rates measured in flight B 14 for members of the ion families $H_3O^+(H_2O)_n$ (●), $H^+X(H_2O)_n$ (■) and $H^+X_2(H_2O)_n$ (▲) as function of the total number of ligands N.

Three out of the minor ions 37 ± 1, 42 ± 1, 83 ± 1 and 114 ± 1 may be readily identified according to their masses as H^+H_2O, $H^+X_1H^+X_2$ and $H^+X(H_2O)_4$ which are members of the main ion families. For 37 ± 1, 83 ± 1 and 114 ± 1 this is supported by chemical arguments as the abundances of these ions fit very well the distributions shown in Fig. 2. The other minor ions (Table 1) do not belong to the major ion families. Like the X ions they are probably formed from proton hydrates by reactions involving trace molecules Y. They may be formed from X-ions if Y molecules have sufficiently large proton affinities. Thus, the reactions involving Y

$$H^+(H_2O)_n + Y \rightarrow H^+Y(H_2O)_{n-1} + H_2O \qquad (1)$$

$$H^+X(H_2O)_i + Y \rightarrow H^+YX(H_2O)_{i-1} + H_2O \qquad (2a)$$

$$\rightarrow H^+Y(H_2O)_i + X \qquad (2b)$$

and the subsequent reaction involving X

$$H^+Y(H_2O)_{n-1} + X \rightarrow H^+YX(H_2O)_{n-2} + H_2O \qquad (3)$$

would give rise to the formation of ions having the general form $H^+Y(H_2O)_n$ and $H^+YX(H_2O)_n$. On their formation the product ions containing Y and possibly also X undergo rapid hydration and individual equilibrium distributions should be established for each ion family.

For ions of the types $H^+Y(H_2O)_n$ and $H^+YX(H_2O)_n$ the distributions should be similar to those for $H^+X(H_2O)_n$ and $H^+X_2(H_2O)_n$ ions (Fig. 2). This is also suggested by our laboratory measurements of cluster ion distributions which involved X and Y molecules as, for example, CH_3CN and CH_3OH (ref. 6). Consequently, if the minor ions listed in Table 2 have the form $H^+Y(H_2O)_n$ or $H^+YX(H_2O)_n$ they would be expected to contain two or three ligands. Species containing one and four ligands should be much less abundant and therefore were probably not detectable.

The total fractional abundances for the families $H^+Y(H_2O)_n$ and $H^+YX(H_2O)_n$ depend on the abundance of Y. A steady-state treatment of the formation of ions containing Y by processes (1) and (2) and of their loss by ion–ion recombination (rate coefficient α) yields

$$\{n(H^+(H_2O)_n + n(H^+X(H_2O)_n)\}kn(Y) = \{n(H^+Y(H_2O)_n) + n(H^+YX(H_2O)_n)\}\alpha n_- \qquad (4)$$

where $n(\ )$ denotes the number density of the species inside the bracket and n_- is the total negative ion concentration. Assuming that the minor ions are of the type $H^+Y(H_2O)_n$ or $H^+YX(H_2O)_n$

and considering that their total fractional abundance is of the order of 0.01, expression (4) yields $n(\mathrm{Y})$ of the order of 1,000 cm^{-3}. Here, k, α and n_- were taken as 2×10^{-9} $cm^3\,s^{-1}$, 10^{-7} $cm^3\,s^{-1}$ and 2,000 cm^{-3} (ref. 5). Thus, Y molecules seem to be ~100 times less abundant than X whose concentration is of the order of 10^5 cm^{-3} (refs 2, 5). Therefore, it is unlikely that ions containing more than one Y were detected. The total abundance of the $H^+YX(H_2O)_n$ family should not exceed that of the $H^+Y(H_2O)_n$ family by much more than a factor 2. Considering an X concentration of the order of 10^5 cm^{-3}, as inferred previously, the time scale for process (3) may be as small as the ion recombination lifetime.

By analogy to the X-ions the Y-ions might be expected to have the form $HY^+(H_2O)_2$ and $HY^+(H_2O)_3$. Mixed ligand clusters having the form $HY^+X(H_2O)_n$ with n being mostly 1 or 2 may also be present in lower fractional abundances. It is also conceivable that instead of a protonated core H^+Y cluster ions contain an atomic core A^+ which has a low ionization potential as, for example, a metal atom. Assuming that the minor ions listed in Table 1 contain two or three ligands which may be H_2O or an X molecule and H_2O molecules mass numbers for the ion cores may be inferred (see Table 1). As the same core may be contained in more than one of the detected ion species the identification of a particular core becomes increasingly likely the more ion species that can be related to it. Unfortunately, however, because some of the minor ions are also sufficiently abundant to be detectable they may be masked by large adjacent mass peaks which complicates the grouping of ion species into families.

Table 2 Tentative identifications for core ions, and proton affinities for protonated molecules

Mass	Core ion	Mass	Proton affinities (kcal mol^{-1})
23 ± 1	Na^+	23	—
27 ± 1	Al^+	27	—
	H^+HCN	28	178.9 ~ 179
32 ± 1	H^+CH_3OH	33	182
	$H^+CH_3NH_2$	32	217
48 ± 1	$H^+CH_2O_2$	47	178
	H^+CH_4S	49	186
	$H^+C_2H_6O$	47	187
63 ± 1	H^+CH_3NO2	62	188
74 ± 1	$H^+C_2H_3NS$	74	187

The minimum number of cores that the 14 minor ion species listed in Table 1 may be related to is 6, including the core masses 23 ± 1, 27 ± 1, 32 ± 1, 48 ± 1, 63 ± 1 and 74 ± 1. However, because there are alternative mass identifications of cores the core masses inferred above are not unique. For example, ions 63 ± 1 and 81 ± 1 which may be hydrates of a core 27 ± 1 and contain two and three water ligands were found in flight B14 to have fractional count rates (relative to the total count rate of this family) of ~25 and 75%. These values are almost the same as those of the corresponding hydrates of H^+X (see Fig. 2) which supports the above identification for 63 ± 1 and 81 ± 1.

The ion composition data previously obtained by rocket-borne mass spectrometers were affected by strong collision-induced cluster ion dissociation and had revealed the presence of an ion 29 ± 2 which has not yet been satisfactorily explained. This ion species may, in fact, be identical to the above ion core 27 ± 1.

Tentative identifications for the above ion cores are given in Table 2. No gaseous compound from which the cores may have formed has so far been detected in the stratosphere. Although CH_3OH and CH_3NO_2 have been considered as stratospheric trace gases[9]. The core ion 27 ± 1 may also be Al^+. However, Al^+ reacts efficiently with O_3 to give AlO^+ (F. Fehsenfeld, personal communication) which makes it an unlikely candidate.

Contaminant gases desorbing from the balloon or the payload may reside for a relatively long period of time in the vicinity of the payload as this moves only at low speed (1–10 cm s^{-1}) relative to the atmospheric gas. Thus, the payload may, in fact, be surrounded by a cloud of contaminant gases containing compounds which have the potential to react with atmospheric ions approaching the payload. In this respect, balloon-borne ion composition measurements differ greatly from rocket-borne measurements in the stratosphere which should not be affected by contamination as relative velocities are of the order of 1 km s^{-1} and thus contaminants should be carried away from the mass spectrometer by the gas flow. Therefore, ions detected in rocket-borne measurements should at least contain the same ion cores as the ambient stratospheric ions. As mentioned before, they may, however, have lost ligand molecules due to collisional dissociation.

The ions 42 ± 2 and 29 ± 2 detected by the rocket measurements should, therefore, represent cores of stratospheric ions. Whether other core ions inferred from the present measurements are also unaffected by contamination cannot be decided.

Future balloon-borne ion composition measurements should therefore be carried out in conditions of rapid motion of the probe relative to the atmospheric gas which reduces the chance that atmospheric ions collide with contaminants. Technically, this can be achieved by using a balloon which descends at a controlled speed of about 1 m s^{-1} which still allows for sufficient integration time.

We thank E. Ferguson and P. Crutzen for helpful discussions and F. Fehsenfeld for information on recent unpublished measurements of Al^+ reactions. Part of this project was funded by the Bundesministerium für Forschung und Technologie, through Gesellschaft für Strahlen-und-Umweltforschung.

Received 2 January; accepted 19 March 1981.

1. Arnold, F., Krankowsky, D. & Marien, K. H. *Nature* **267**, 30 (1977).
2. Arnold, F., Böhringer, H. & Henschen, G. *Geophys. Res. Lett.* **5**, 653 (1978).
3. Arijs, E., Ingels, J. & Nevegans, D. *Nature* **271**, 642 (1978).
4. Olsen, J. B., Amme, R. C., Brooks, J. N., Murcray, D. G. & Keller, G. E. *Trans. Am. geophys. Un.* **58**, 1201 (1977).
5. Arnold, F., Henschen, G. & Ferguson, E. E. *Planet. Space Sci.* (in the press).
6. Böhringer, H. & Arnold, F. *Nature* (in the press).
7. Kebarle, P., Haynes, R. N. & Collins, J. G. *J. Am. chem. Soc.* **89**, 5753 (1967).
8. Harries, J. E. *Rev. Geophys. Space Phys.* **14**, 565 (1976).
9. Shimazaki, T. & Ogawa, T. *J. geophys. Res.* **79**, 3411 (1974).
10. Fehsenfeld, F., Dotan, F. C., Albritton, D., Howard, C. & Ferguson, E. *J. geophys. Res.* **83**, 1333 (1978).

Carbon dioxide, ammonia and the origin of life

T. M. L. Wigley* & P. Brimblecombe†

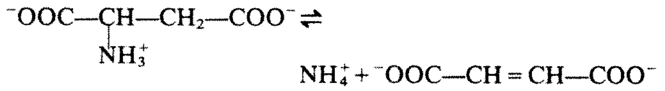
Climatic Research Unit*, and School of Environmental Sciences†,
University of East Anglia, Norwich NR4 7TJ, UK

Stellar evolution theory predicts that the luminosity of the Sun has increased by ~30% over the past 4,000 Myr. Yet geological and biological evidence indicates that the climate of the Earth between 3,000 and 4,000 Myr ago was as warm as, or warmer than, today. This apparent contradiction, the 'faint Sun paradox', has been resolved by invoking the greenhouse effect of radiatively active gases in the early Earth atmosphere. Sagan and Mullen[1] first suggested that the concentration of ammonia in the early atmosphere was around 10–100 p.p.m., sufficiently high to counteract the reduced luminosity. However, because ammonia photodissociates readily and has a short atmospheric residence time[2,3], such a concentration could be maintained only by a large continuous ammonia source. For this reason, carbon dioxide is now considered to have been the radiatively active gas[4–7]. Some atmospheric ammonia is, nevertheless, required to provide conditions conducive to the origin of life[8]. We now show that, if the early Earth's atmosphere contained high concentrations of CO_2, as suggested above, then the chemical conditions required for life to begin can be maintained by very low ammonia partial pressures, rather similar to those observed today.

0028–0836/81/210213—03$01.00

The conditions believed necessary for life to have evolved on Earth are: the existence of liquid water; and aqueous chemical conditions which permit a stable solution of simple amino acids, the chemical precursors of life. Organic molecules may form in the absence of ammonia, but ammonia is probably essential to ensure the stability of amino acids. Amino acid stability depends on temperature, pH and the concentration of dissolved NH_3 (ref. 8). The early evolution of life, therefore, places constraints on the chemistry of early aqueous environments, and hence on the atmospheric concentration of ammonia. To quantify these constraints, we will consider two end-member aqueous systems, a seawater and a freshwater, as possible life-origin environments. First, however, we will determine a general lower bound to aqueous ammonium ion activity as a function of pH and P_{NH_3}.

A minimum ammonium ion activity is required to prevent the decomposition by reversible deamination of amino acids to the corresponding hydroxy acids. Bada and Miller[8] consider the deamination of aspartic acid to fumaric acid



$$^{-}OOC{-}\underset{NH_3^+}{CH}{-}CH_2{-}COO^{-} \rightleftharpoons NH_4^+ + {}^{-}OOC{-}CH{=}CH{-}COO^{-}$$

$$pK = -8.22 + 2.276/T + 0.0106T;\ pX \equiv -\log_{10} X,\ T \text{ in K} \quad (1)$$

and suggest that a necessary condition for life to begin is that the ratio of total aspartate to total fumarate should not fall much below one. At near-neutral pH the aspartate-fumarate ratio sets a lower limit to the ammonium ion activity of around $10^{-2.6}$ (at 25 °C). At lower pH values, higher ammonium ion concentrations are required to offset the higher stability of fumarate to protonation. From equation (1) and the acidity constants for aspartic and fumaric acids we have

$$\text{Asp}^{\text{TOT}}/\text{Fum}^{\text{TOT}} \approx a\text{NH}_4^+\, 10^{-pK}\alpha_a/\alpha_f \quad (2)$$

Here

$$\alpha_a = \frac{K_a'''}{a\text{H}^+} + 1 + \frac{a\text{H}^+}{K_a''} + \frac{(a\text{H}^+)^2}{K_a'} \quad \text{and} \quad \alpha_f = 1 + \frac{a\text{H}^+}{K_f''} + \frac{(a\text{H}^+)^2}{K_f'}$$

where K_a''', K_a'' and K_a' are the dissociation constants for aspartic acid, K_f'' and K_f' those for fumaric acid, and $a\text{H}^+$ ($\equiv 10^{-pH}$) is the hydrogen ion activity. At 25 °C, $pK_a''' \approx 9.9$, $pK_a'' \approx 3.8$, $pK_a' \approx 1.8$, $pK_f'' \approx 4.4$ and $pK_f' \approx 7.4$. Equation (2) is obtained from the acidity reactions of aspartic and fumaric acids neglecting the effect of ionic strength on the activities of the organic acid species. The stability constants quoted are from ref. 10. For $5.3 \leqslant pH \leqslant 8.4$, $\alpha_a \approx 1$, and for $pH \geqslant 5.4$, $\alpha_f \approx 1$. The isopleth $\text{Asp}^{\text{TOT}}/\text{Fum}^{\text{TOT}} = 1$ (at 25 °C), curve (1) on Fig. 1, places an approximate lower limit on the ammonium ion activity for life to evolve.

Equation (2) and the constraint $\text{Asp}^{\text{TOT}}/\text{Fum}^{\text{TOT}} = 1$ determine a minimum ammonium ion activity as a function of pH. The corresponding ammonia partial pressures are determined by the dissociation reaction of NH_4^+. From the mass action equation for this reaction we have

$$pP_{NH_3} = pK_a - pK_{H,a} + p(a\text{NH}_4^+) - p\text{H} \quad (3)$$

where $K_{H,a}$ is the Henry's law constant for ammonia ($pK_{H,a} = -1.785 - 34.85(298.15 - T)/5.7118\ T$, based on free energy and enthalpy data in ref. 11) and K_a is the dissociation constant for NH_4^+ ($pK_a = 0.6322 - 0.001225\ T - 2835.76/T$; ref. 12). Lines of constant ammonia partial pressure at 25 °C based on this equation are shown in Fig. 1.

We now consider the case where life began in a seawater environment in which chemical conditions were controlled by large-scale factors. In these circumstances, an extreme upper limit to the ammonium ion concentration may be set by dissolving all the available atmospheric nitrogen in the oceans as ammonia[8] (curve (2) in Fig. 1). A more realistic upper limit has been suggested[8] based on the cation exchange properties of clay minerals (curve (3) in Fig. 1). The zone between curves (1) and (3) in Fig. 1 defines the range of suitable ammonia partial pressures for life to evolve. pH is the main controlling factor. As the early oceans were probably much more acidic than today[13], extremely low ammonia partial pressures (as low as 10^{-8} atm at pH 5.6, see Fig. 1) would be sufficient to ensure a ratio of aspartate to fumarate in excess of 1. Present-day values, while subject to some uncertainty due to measurement difficulties, are of the order of $2\text{–}6 \times 10^{-8}$ atm over ocean areas[14,15].

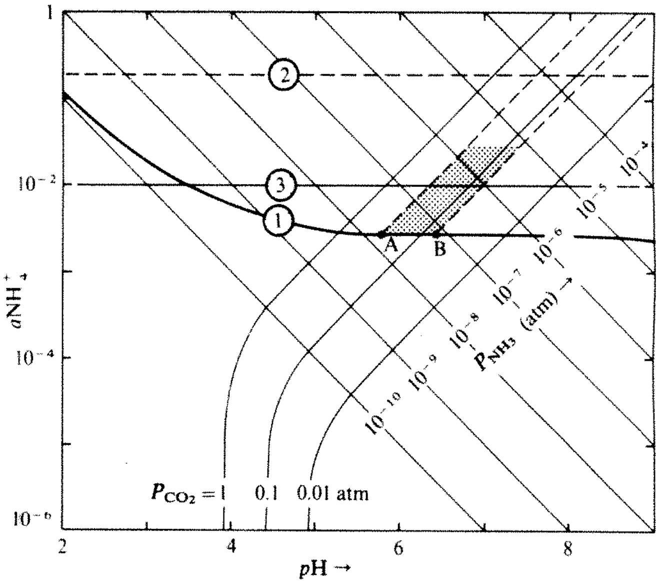


Fig. 1 Curve (1) and lines (2) and (3) define approximate lower (1) and upper (2 or 3) boundaries for ammonium ion activity as a function of pH for the amino acid stability required to allow life to evolve. The ordinate gives ammonium ion activity on a log scale. Isopleths of constant P_{NH_3} are shown as diagonal lines sloping from left to right calculated using equation (3). The zone between curves (1) and (3) is that which would be suitable for life to begin in an oceanic environment. At moderate pH ($\leqslant 6$), suitable conditions would prevail even for low P_{NH_3} ($\leqslant 10^{-8}$ atm). In a freshwater environment, chemical conditions are additionally influenced by P_{CO_2}. Curves of constant P_{CO_2} are also shown (based on equation (8)). These relate the four variables, P_{CO_2}, P_{NH_3}, $a\text{NH}_4^+$ and pH in freshwater environments. A 'life zone' (stippled region) is defined by the range of P_{CO_2} partial pressures thought to have prevailed between 3,500 and 4,250 Myr BP (ref. 20), and by the amino acid stability curve (1). This figure is based on thermodynamic data for 25 °C.

There is, however, no *a priori* reason to assume that life originated in the early oceans. We must, therefore, consider the small-scale end-member possibility, a local 'fresh' water environment. Here the high atmospheric carbon dioxide partial pressures thought to have prevailed at the time are likely to have controlled pH values. To evaluate the effect of high P_{CO_2} on pH we consider a simple model in which only ammonium and carbonate species are present. Charge balance necessitates that

$$m\text{NH}_4^+ + m\text{H}^+ = m\text{OH}^- + m\text{HCO}_3^- + 2m\text{CO}_3^= \quad (4)$$

where m denotes molality. For $P_{CO_2} \geqslant 10^{-5}$ atm, $pH \leqslant 9$ and $P_{NH_3} \geqslant 10^{-10}$ atm, ammonium and bicarbonate ions are the dominant species and the charge balance equation becomes

$$m\text{NH}_4^+ = m\text{HCO}_3^- \quad (5)$$

Within the accuracy of this result molalities may be replaced by activities, as the activity coefficients of NH_4^+ and HCO_3^-, as determined by the extended Debye-Hückel formulation, are approximately equal. Hence

$$a\text{NH}_4^+ = P_{CO_2} K_{H,c} K_c / a\text{H}^+ \quad (6)$$

This result comes directly from equation (5) using $a\text{H}_2\text{CO}_3^* = K_{H,c} P_{CO_2}$ where $pK_{H,c} = 14.0184 - 0.015264\ T - 2385.73/T$ is the Henry's Law constant for carbon dioxide[16], and the mass

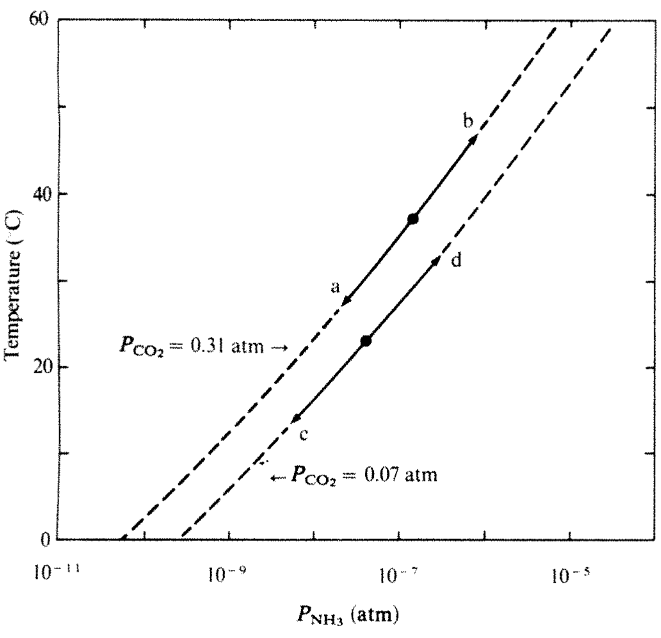


Fig. 2 The effect of changing temperature and P_{CO_2} on the minimum ammonia partial pressure needed for life to evolve. The points a and b span conditions which are thought to have prevailed around 4,250 Myr BP, while c and d span conditions for 3,500 Myr BP. If temperature is lower, then lower P_{NH_3} values are sufficient. The coupling of temperature and P_{CO_2} through the greenhouse effect means that life could have originated at lower P_{NH_3} for lower P_{CO_2} (see points a and c), in spite of the fact that the direct influence of P_{CO_2} is in the opposite direction.

action equation for the dissociation of $H_2CO_3^*$, $aHCO_3^- \, aH^+ = K_c \, aH_2CO_3^*$ where $pK_c = -14.8435 + 0.032786\,T + 3404.71/T$ (ref. 16). $H_2CO_3^*$ is the sum of dissolved carbon dioxide and carbonic acid. Equation (6) can be written in logarithmic form as

$$pP_{CO_2} = -pK_c - pK_{H,c} + p(aNH_4^+) + pH \quad (7)$$

which allows isopleths of constant carbon dioxide partial pressure to be constructed on Fig. 1. Strictly this linear relationship is only valid for $P_{NH_3} \geqslant 10^{-10}$ atm. Although lower ammonia partial pressures are not relevant here the more general relationship, obtainable from equation (4),

$$10^{-pH} + aNH_4^+ \approx 10^{pH}[10^{-14} + 10^{-7.8} P_{CO_2}(1 + 10^{-10+pH})] \quad (8)$$

has also been used in constructing Fig. 1. In deriving equation (8) activity coefficients have been neglected (compare with derivation of equation (7)), but this results in no significant loss of accuracy.

Life is believed to have originated within the Earth's first billion years, possibly earlier than 3,000 Myr ago (ref. 17; but see also refs 18 and 19). Hart[20] estimates that the atmospheric P_{CO_2} was 0.07 atm at 3,500 Myr BP and 0.31 atm at 4,250 Myr BP, and values of this magnitude are required to give a sufficient greenhouse effect to compensate for reduced solar luminosity[4–7]. The range 0.07–0.31 atm P_{CO_2} therefore determines the range of pH values in early freshwater environments. Combined with the lower limit to aNH_4^+ derived earlier this range defines a 'life zone' in Fig. 1 (stippled region) within which chemical conditions would have been conducive to the origin of life. Note that the upper limit imposed by cation exchange considerations (curve (3) in Fig. 1) need not apply to the freshwater environment.

Figure 1 shows immediately that for life to evolve the local P_{NH_3} pressure must exceed 10^{-8} atm. Present-day values in non-oceanic areas are usually less than this value, but may exceed it. Values quoted recently range between 0.1 and 1.2 10^{-8} atm. (refs 14 and 21), and even higher values have been reported[22]. All these data may reflect local organic processes and so are not strictly relevant to conditions which might have prevailed before the origin of life. However, similar levels could easily be attained by the catalytic photochemical reduction of atmospheric nitrogen by TiO_2 in desert sands. Henderson-Sellers and Schwartz[23] postulate a 100-fold increase above present productivity by this mechanism to estimate a total atmospheric ammonia loading of 10^{13} kg, which corresponds to a partial pressure of around 1.9×10^{-6} atm. Using their figures, and present productivity levels, gives an estimated global atmospheric ammonia mass of 7×10^{10}–7×10^{12} kg, equivalent to a partial pressure between 1.3×10^{-8} and 1.3×10^{-6} atm. While these values may be a little high compared with present-day measurements[14,21] the lower end seems reasonable for local environments.

Figure 1 only shows the results for 25 °C, but we have performed the same calculations at other temperatures. The effect of temperature is shown in Fig. 2 where the ammonia partial pressure minimum spanned by the freshwater 'life zone' in Fig. 1 (points A and B) is plotted as a function of temperature. As temperature decreases the minimum ammonia partial pressure decreases. Mean global temperature estimates corresponding to the limiting P_{CO_2} values of 0.31 and 0.07 atm assumed above have been given by Owen *et al.*[6]. Assuming a global temperature range of ± 10 °C we can define regions in Fig. 2 which correspond to the range of P_{CO_2} and temperature values likely for the early Earth. If life began in a low salinity aqueous environment (such as might exist where local conditions today are conducive to the inorganic production of ammonia) then local P_{NH_3} values around 10^{-7} atm and possibly as low as 10^{-8} would would have been sufficient to have maintained a suitable ratio of amino to hydroxy acids. Global mean P_{NH_3} may well have been even smaller.

Life may have been originated in highly saline conditions; the seawater case considered initially. In this case low ammonia partial pressures ($\leqslant 10^{-8}$ atm) are still sufficient, provided the pH is relatively low ($p\text{H} \leqslant 6$) compared with modern oceanic environments. Thus, high partial pressures of ammonia are not essential precursors for the origin of life in either of the systems considered here. Although our P_{NH_3} estimates are substantially less than some other estimates ($\sim 10^{-4.5}$ atm, see refs 1, 20, 24), and much less than the limit ($\sim 10^{-5}$ atm) for which a noticeable greenhouse effect might occur[3,23], such values seem unlikely because of the rapidity with which ammonia is photodissociated[2,3]. Our values are similar to values which occur locally today. A strongly reducing atmosphere does not seem to be necessary for the origin of life.

Received 4 February; accepted 1 April 1981.

1. Sagan, C. & Mullen, G. *Science* **177,** 52 (1972).
2. Abelson, P. H. *Proc. natn. Acad. Sci. U.S.A.* **55,** 1365 (1966).
3. Kuhn, W. R. & Atreya, S. K. *Icarus* **37,** 207 (1979).
4. Owen, T. in *Evolution of Planetary Atmospheres and Climatology of the Earth*, 1–10 (CNES, Toulouse, 1978).
5. Henderson-Sellers, A. & Meadows, A. J. in *Evolution of Planetary Atmospheres and Climatology of the Earth*, 25–30 (CNES, Toulouse, 1978).
6. Owen, T., Cess, R. D. & Ramanathan, V. *Nature* **277,** 640 (1979).
7. Henderson-Sellers, A., Benlow, A. & Meadows, A. J. *Q. Jl R. astr. Soc.* **21,** 74 (1980).
8. Bada, J. L. & Miller, S. L. *Science* **159,** 423 (1968).
9. Pinto, J. P., Gladstone, G. R. & Yung, Y. L. *Science* **210,** 183 (1980).
10. Sillén, L. G. & Martell, A. E. *Stability Constants of Metal-Ion Complexes* (Chemical Society Special Publication No. 17, 1964).
11. Robie, R. A., Hemingway, B. S. & Fisher, J. R. *Bull. geol. Surv.* 1452 (1978).
12. Wright, J. M., Lindsay, W. T. Jr & Druga, T. R. *U. S. Atomic Energy Commission R and D Rep. WAPD-TM*-204 (1961).
13. Rubey, W. W. *Bull. geol. Soc. Am.* **62,** 1111 (1951).
14. Söderlund, R. & Svensson, B. H. in *Nitrogen, Phosphorus and Sulphur: Global Cycles* (eds. Svensson, B. H. & Söderlund, R) 23–73 (SCOPE Rep. 7, Stockholm, 1976).
15. Georgii, H. W. & Gravenhorst, G. *Pure appl. Geophys.* **115,** 503 (1977).
16. Harned, H. S. & Davies, R. Jr *J. Am. chem. Soc.* **65,** 2030 (1943).
17. Schidlowski, N., Appel, P. W. V., Eichmann, R. & Junge, C. E. *Geochim. cosmochim. Acta* **43,** 189 (1979).
18. Bridgwater, D. *et al. Nature* **289,** 51 (1981).
19. Nagy, B., Engel, M. H., Zumberge, J. E., Ogino, H. & Chang, S. Y. *Nature* **289,** 53 (1981).
20. Hart, M. H. *Icarus* **33,** 23 (1978).
21. Levine, J. S., Augustsson, T. R. & Hoell, J. M. *Geophys. Res. Lett.* **7,** 317 (1980).
22. Lodge, J. P., Machado, P. A., Pate, J. B., Sheesly, D. C. Wartburg, A. F. *Tellus* **26,** 250 (1974).
23. Henderson-Sellers, A. & Schwartz, A. W. *Nature* **287,** 526 (1980).
24. Sagan, C. *Nature* **269,** 224 (1977).

Archaeomagnetic secular variation record of Mount Vesuvius

G. S. Hoye

Institute of Earth and Planetary Physics, University of Alberta, Edmonton, Alberta, Canada T6G 2J1

The lavas produced by historically recorded eruptions of Mt Vesuvius, and the Herculaneum mudflow resulting from the AD 79 eruption have been investigated archaeomagnetically. The mudflow carries a magnetically hard remanence, the direction of which is significantly different from that of the present-day field, suggesting that the remanence is primary. The secular variation curve for Vesuvius agrees well with the Rome record and with other curves from Italy and London. A model for westward drift which accounts for certain features of secular variation in Europe over the past 200 yr requires a mean deceleration rate of 0.004 deg yr^{-2}. This is less than the current rate, suggesting that the deceleration rate was smaller 200 yr ago than it is today.

The historical lava-producing eruptions of Mt Vesuvius between AD 1631 and 1944 have recently been mapped in detail[1]. Earlier events are poorly documented except the Plinian eruption of AD 79 which produced no significant lava. The period covered by the flows and the proximity of the volcano to Rome, where direct compass observations have been recorded since the sixteenth century, make an archaeomagnetic investigation particularly attractive because it is possible to compare the observatory with the archaeomagnetic records.

Eleven dated lava flows were sampled: where exposure permitted, flows were sampled at two or more sites (Table 1) located to obtain the greatest possible coverage of each flow. Five orientated 25-mm cores were collected at each site using routine techniques. Twenty orientated specimens of the mudflow which engulfed Herculaneum during the AD 79 eruption were collected from inside two adjacent rooms that had been filled completely by the mud. The finest silty material exposed was sampled by pressing small plastic boxes into prepared vertical surfaces.

Routine alternating field demagnetization indicated that the remanence at all sites was essentially single-component and stable. Angular shifts between demagnetization steps were small and random, and the mean median destructive field was close to 200 Oe. The natural remanent magnetization (NRM) and the cleaned flow mean directions are not significantly different (Table 1), but in all cases the precision is improved (average α_{95} is 2.7° and 1.9° respectively) and consequently the cleaned results are preferred.

The cleaned directions are plotted in Fig. 1*a*. The early part of the curve (excluding the point for AD 79) is indicated by a dashed line because this 330-yr span is defined by only two points. The general features of the curve are: a shift from shallower inclinations between the fourteenth and seventeenth centuries, and an incomplete clockwise elliptical loop in the subsequent period up to the present.

Observations during the past few centuries show that the curve of secular variation in Europe has been proceeding in a near-elliptical clockwise loop; Fig. 1*b* shows the curve for Rome. The shapes of the curves in Fig. 1*a* and *b* agree, bearing in mind that the 95% confidence limits on the Vesuvian directions are typically 1° or 2°. Inclinations at Vesuvius are expected to be ~1° shallower due to the latitude difference, but even after allowing for this, the Vesuvian curve remains 2° or 3° too shallow. A similar discrepancy was noted in the lavas of Mt Etna by Tanguy[2] who suggested that it was caused by magnetic refraction. Tanguy's results are replotted in Fig. 1*c*. The latitude difference requires that this curve be shifted ~3° to steeper inclinations for comparison with the Vesuvian results, and when this is done the elliptical portions of the two curves show good agreement. There is, however, a significant difference in the early parts of the curves: Tanguy's curve seems to describe part of a counterclockwise loop with no indication of movement from shallower inclinations. Mt Etna is only 300 km south of Ischia and it seems unlikely that the difference is due to real variations in the non-dipole field. The discrepancy does depend to a large extent on the point for the lava flow of 1329. If this flow were wrongly identified (the incompleteness of early records makes this a possibility) the point may be disregarded, and this brings the two curves into better agreement over a longer interval. Tanguy expressed doubt about the reliability of the date.

It is interesting to compare the present observations with the record for London[3] which displays a large shift from shallower inclinations between AD 1300 and 1600 followed by the well-documented clockwise elliptical loop. The corresponding movement in Italy is not so large, but identical angular displacements are not to be expected because the two localities are

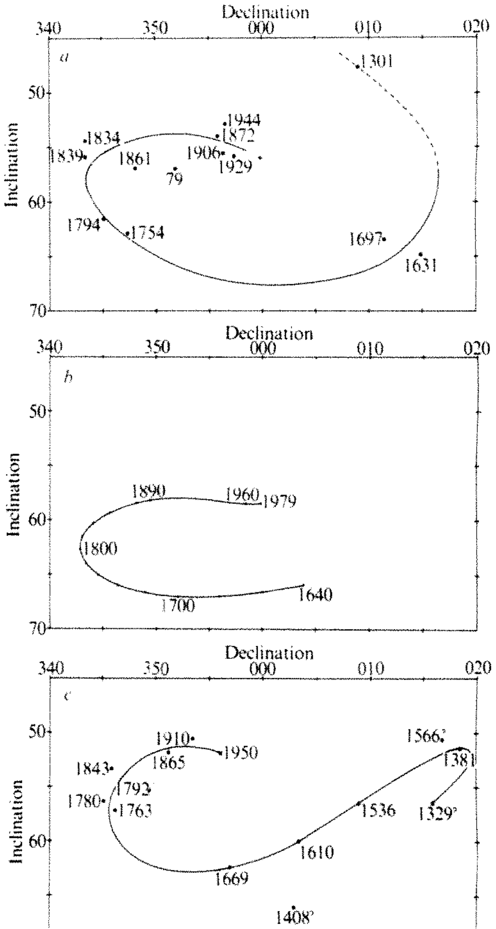


Fig. 1 *a*, The secular variation record for Mt Vesuvius. The dated points show the cleaned flow mean directions in Table 1. The present field direction is indicated by the cross. The direction for AD 1301 is from the lava del' Arso, Ischia[6] which is only 40 km to the west of Vesuvius. The point for AD 79 is obtained from the mudflow at Herculaneum. *b*, The secular variation record for Rome derived from observatory records. 1979–60 supplied by F. Molina (personal communication); 1890–1640 taken from ref. 7. Bauer's reconstruction for the period 1640–1508 should be disregarded although the easterly declination is not in doubt. Bauer describes an observation of declination 005° for 1436, and there is a record of declination 007° for the eastern Mediterranean in the mid-sixteenth century[8]. *c*, The secular variation record for Mt Etna from Tanguy[2]. A question mark by a point indicates that he doubted the date of the flow sampled. The point for 1950 is from ref. 9.

0028–0836/81/210216—03$01.00

Table 1 Archaeomagnetic results from Vesuvius and from the lava del' Arso, Ischia*

Date of flow†	Site	*J*	Dec	Inc	α_{95}	*k*	*N*
1944	A	4.61	353.2	53.6	1.6	2,381	5
			353.0	53.6	1.8	1,884	
	B	4.05	359.7	50.7	1.1	4,871	5
			000.1	52.1	1.1	4,578	
		4.33	**356.5**	**52.2**	**1.7**	**771**	**10**
			356.6	**52.9**	**1.7**	**833**	
1929	A	4.27	357.1	56.7	1.4	2,988	5
			357.6	56.7	1.2	3,809	
	B	5.03	357.3	54.8	1.5	2,490	5
			357.3	55.2	1.5	2,740	
		4.65	**357.2**	**55.7**	**1.1**	**2,103**	**10**
			357.4	**55.9**	**0.9**	**2,693**	
1906		**4.88**	**356.1**	**55.8**	**1.2**	**3,838**	**5**
			356.5	**55.6**	**1.1**	**4,842**	
1872	A	4.47	355.8	56.1	4.2	271	5
			357.8	55.3	2.4	988	
	B	4.22	358.1	49.7	4.7	271	5
			354.0	52.6	1.9	1,633	
		4.35	**357.0**	**52.9**	**3.3**	**209**	**10**
			355.8	**54.0**	**1.7**	**807**	
1861		**3.57**	**348.1**	**54.3**	**5.0**	**237**	**5**
			348.2	**57.0**	**3.1**	**593**	
1839		**3.53**	**343.3**	**55.7**	**2.7**	**819**	**5**
			343.4	**55.8**	**2.3**	**1,132**	
1834	A	4.05	341.5	53.5	3.2	580	5
			340.9	53.1	3.0	645	
	B	3.82	347.3	54.7	2.6	861	5
			346.0	55.5	1.8	1,882	
		3.94	**344.4**	**54.2**	**2.1**	**550**	**10**
			343.4	**54.4**	**1.9**	**653**	
1794		**3.98**	**345.7**	**60.4**	**1.7**	**1,928**	**5**
			345.2	**61.6**	**1.0**	**5,614**	
1754	A	4.82	347.9	61.9	2.0	1,529	5
			348.3	62.5	2.1	1,335	
	B	6.02	347.2	61.5	3.3	530	5
			347.5	62.9	1.9	1,620	
	C	4.15	346.2	63.3	1.8	2,702	4
			346.2	63.3	1.7	2,930	
		5.00	**347.2**	**62.1**	**1.2**	**1,022**	**14**
			347.4	**62.9**	**0.9**	**1,857**	
1697	A	9.22	006.2	62.4	1.9	1,708	5
			006.2	62.6	1.7	2,017	
	B	6.29	018.7	63.5	2.4	1,000	5
			017.2	64.2	1.9	1,611	
		7.76	**012.3**	**63.1**	**2.2**	**471**	**10**
			011.5	**63.5**	**1.9**	**618**	
1631	A	3.14	020.2	57.3	7.1	118	5
			014.2	63.4	1.5	2,553	
	B	4.15	014.4	64.9	1.9	1,682	5
			015.6	66.2	3.1	627	
		3.65	**017.7**	**61.2**	**4.0**	**147**	**10**
			014.9	**64.7**	**1.7**	**772**	
1301*		**3.2**	**010.6**	**47.3**	**2.1**	**209**	**24**
			009.0	**47.7**	**1.2**	**566**	**14**
79‡		**0.28**	**350.2**	**54.9**	**5.3**	**39**	**20**
			351.9	**57.0**	**2.2**	**212**	

Entries in normal type refer to sites. Those in bold type give flow mean data. The first line of each entry (site or flow) refers to the NRM, the second to the partially demagnetized remanence. In cases where there is only one site per flow, site and flow means are the same. *J*, site or flow mean intensity of magnetization in 10^{-3} e.m.u. cm^{-3}; Dec, declination; Inc, inclination; α_{95}, Fisherian 95% confidence cone about the mean; *k*, Fisherian precision parameter; *N*, number of individually orientated specimens used for the statistics. One core from a site in the flow of 1754 was misorientated and its direction rejected.

* Taken from ref. 6. † Dates are all AD and are taken from the map of Principe[1]. Flows shown on the available topographical maps are not always reliably identified. ‡ Mudflow.

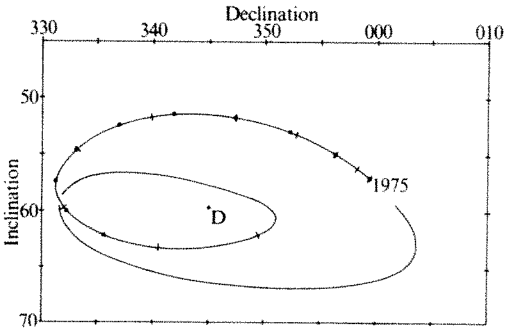


Fig. 2 Two models for the secular variation curve at Vesuvius, based on the IGRF75 representation of the main field. The dots show the predicted direction at 100-yr intervals from 1975, for a uniform drift rate of 0.1 deg yr^{-1}. The transverse bars give the direction at 25-yr intervals from 1975, for a deceleration rate of 0.004 deg yr^{-2} to a velocity of 0.1 deg yr^{-1} in 1975. Point D indicates the dipole field direction.

far enough apart to be affected in different amounts by the same non-dipole anomalies.

It is not possible to extend the secular variation curve to include the point for Herculaneum because of the interval of 1,200 yr between this point and that for Ischia. The direction for Herculaneum, which lies within the recent ellipse, is not unreasonable, however. In the London record, the direction for AD 79 lies just outside the ellipse. This similarity in position, and the stability of the remanence, suggest that the mudflow did indeed record the ambient field direction at the time of emplacement. The mechanism of magnetization remains uninvestigated but it may be a shear-induced remanence[4] acquired when the mud came to rest.

A theoretical secular variation curve for Vesuvius based on the 1975 International Geomagnetic Reference Field (IGRF75) is illustrated in Fig. 2. The direction of the dipole component of the main field, which is defined by the first three spherical harmonic coefficients, has remained effectively constant since 1829 whereas the dipole moment has decreased by 6%. In the model the dipole contribution is taken as fixed, and the non-dipole field, which is given by the higher-order coefficients, is driven westwards at the rate of 0.1 deg yr^{-1}. This rate is a reasonable estimate for 1975 and is taken from ref. 5. The secular variation curve for this model consists of a double clockwise loop about the dipole direction (Fig. 2). Although the whole curve is shifted to the west because of the orientation of the stationary dipole, the ranges of declination and inclination and the general elliptical shape are comparable with observation (Fig. 1). However, the time scale, which is determined by the rate of westward drift, is not suitable. For example, the time of maximum westward declination which was ~1810 occurs in the model ~1250. In fact the model is in fair agreement with observation only in the present century. Better agreement over the past two centuries can be obtained if the westward drift is allowed to decelerate. A first-order calculation for a uniform deceleration which brings the time of maximum westward declination to ~1810 (Fig. 2) yields a mean deceleration of 0.004 deg yr^{-2}. This is about half the observed rate, averaged over the past three decades of 0.01 deg yr^{-2} (derived from Fig. 2 of ref. 5). The same calculation for London yields the same mean deceleration, which suggests that westward deceleration of the non-dipole field has been occurring on a continental scale in Europe over the past two centuries. The absence of a double loop in the record implies that the model breaks down after a few centuries, and this is in accordance with the recognized transient behaviour of non-dipole features.

This work was supported by the Natural Sciences and Engineering Research Council of Canada and by the President's Fund of the University of Alberta. The samples from Herculaneum were collected under permit No. 4239 issued by the

Soprintendenza Archeologica di Napoli e Caserta. I thank M. E. Evans for fruitful discussions and R. Scandone for assistance in the field.

Received 2 January; accepted 6 March 1981.

1. Principe, C. thesis, Univ. Pisa (1979).
2. Tanguy, J. C. *Archaeometry* **12,** 115–128 (1970).
3. Aitken, M. J. *Physics and Archaeology* 2nd edn (Oxford University Press, 1974).
4. Games, K. P. *Geophys. J. R. astr. Soc.* **48,** 315–329 (1977).
5. Harwood, J. M. & Malin, S. C. R. *Nature* **259,** 469–471 (1976).
6. Tanguy, J-C. & Pozzi, J-P. *C. r. hebd. Séanc. Acad. Sci. Paris* **274,** 352–354 (1972).
7. Bauer, L. A. *Phys. Rev.* **3,** 34–48 (1899).
8. Brice, W., Imber, C., Lorch, R. & Pelham, P. *Archaeometry* **18,** 228–229 (1976).
9. Lamoureux, C. & Klerkx, J. *Ann. Soc. Geol. Belg.* **90,** 261–277 (1967).

Stromatolites at ~3,500 Myr and a greenstone–granite unconformity in the Zimbabwean Archaean

J. L. Orpen & J. F. Wilson

Geology Department, University of Zimbabwe, PO Box MP 167,
Mount Pleasant, Salisbury, Zimbabwe

Two controversial areas of geological endeavour are the establishment of the antiquity of life and the tectonic setting of greenstone sequences. We record here the recent discoveries in the Fort Victoria greenstone belt of stromatolites in limestones assigned to ~3,500 Myr (minimum age) Sebakwian Group rocks of the Rhodesian Archaean Craton within Zimbabwe, and a nearby outcrop of a thin sedimentary formation, basal to a thick ~2,700 Myr volcanic pile, resting with definite unconformity on ~3,500 Myr Mushandike Granite.

Granite–greenstone terranes of different ages are now known to exist in the Rhodesian Craton[1–3]. The oldest greenstone belt remnants constitute the Sebakwian Group dated at ~3,500 Myr minimum on the evidence from various granites and gneisses; the more extensive main greenstone belts comprise the Bulawayan and Shamvaian Groups (Table 1). An unconformity in the Bulawayan Group allows its subdivision into the Lower Greenstones (?~2,900 Myr) and the Upper Greenstones, which together with the overlying Shamvaian are ~2,700 Myr old. Two suites of late granite post-date greenstone belts. These are the largely tonalitic Sesombi Suite at ~2,700 Myr, and the dominantly adamellitic Chilimanzi Suite at ~2,600 Myr. Archaean stromatolites in Zimbabwe have been recorded only from ~2,700 Myr Upper Greenstones[4]; further the only recorded and clearly demonstrable occurrence of greenstone belt rocks deposited unconformably on granitic basement is from the eastern margin of the Belingwe belt[5].

Table 1 Simplified stratigraphic table of the Zimbabwean Archaean

		Age (Myr)
Chilimanzi Suite (Adamellite)		~2,600
Sesombi Suite (Tonalite)		~2,700
Shamvaian Group		~2,700
Bulawayan Group	Upper Greenstones	~2,700
	Lower Greenstones	?~2,900
Mushandike Granite		~3,500
Sebakwian Group		~3,500

The newly discovered stromatolites occur in limestones some 9 km east of Mushandike Dam wall, in the northern margin of the Fort Victoria greenstone belt, and within the Mushandike National Park (Fig. 1). These limestones are confined to the vicinity of the lake and to the east. They are part of a sedimentary unit that has its greatest development west of Mushandike Dam. This unit was originally regarded by Wilson[6] as belonging to the Bulawayan succession. Later, as part of a craton-wide reassessment, Wilson[2] assigned the bulk of this sedimentary unit and some associated metavolcanic rocks to the ~3,500 Myr Sebakwian Group of the area. He was undecided, however, on the status of the sedimentary rocks east of the Dam because of the effects of later (? ~2,700 Myr) tectonism.

Our recent fieldwork has indicated the presence of two greenstone belt successions in this area (Fig. 1)—one clearly younger, and one which we regard as older than the Mushandike Granite dated at $3{,}445 \pm 260$ Myr ($\lambda\,^{87}\mathrm{Rb} = 1.42 \times 10^{-11}\ \mathrm{yr}^{-1}$; errors at 2σ)[7], equating with the Upper Greenstones and Sebakwian Group respectively.

We propose to call the basal sediments of the Upper Greenstones the Gwenya formation, after Gwenya hill on the western flanks of which an 80-m thick sequence of ferruginous siltstones, with minor arkose and quartzite lenses, lies directly on the Mushandike Granite (Fig. 2). The sediments all dip at 50° SE and young to the south, away from the granite. The unconformity contact is essentially planar with minor undulatory features which may be wash-out channels. It is not faulted.

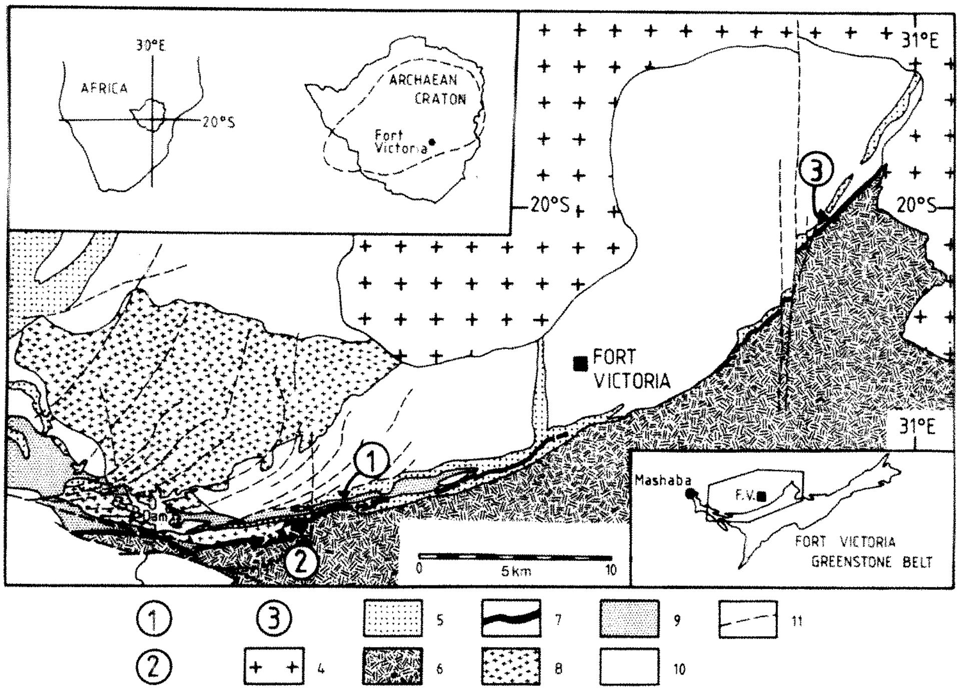


Fig. 1 Simplified geological map of the northern margin of the Fort Victoria greenstone belt. 1, Mushandike stromatolite locality; 2, Gwenya hill unconformity exposure; 3, Rockwood farm unconformity exposure; 4, Chilimanzi granite (~2,600 Myr); 5, Mashaba Igneous Complex; 6, Upper Greenstones volcanic rocks; 7, Gwenya formation (~2,700 Myr); 8, Mushandike Granite (~3,500 Myr); 9, Sebakwian Group; 10, various gneisses; 11, shear zones.

0028-0836/81/210218—03$01.00

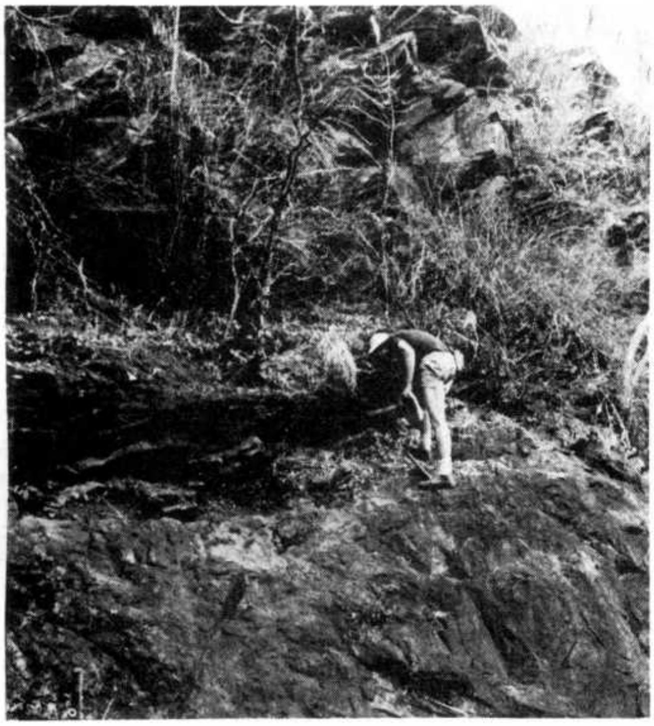


Fig. 2 Gwenya hill unconformity exposure. Left shoe straddles contact between weathered Mushandike Granite regolith and well-cleaved siltstones of Gwenya formation. A shallow depression in the centre of the picture forms a possible wash-out channel now filled with bedded siliceous siltstone.

West of the hill, the Gwenya formation is displayed by minor north–south faults, but seems traceable to south of the Mushandike Dam wall. Dips steepen westwards to 70°–80° and as the wall is approached intense east–west shearing, slightly discordant to strike, has so disrupted the geology that we cannot establish field relationships between granite and sediments. However, Coward and Shackleton[8] claim to have located an unconformity exposure, apparently from this region.

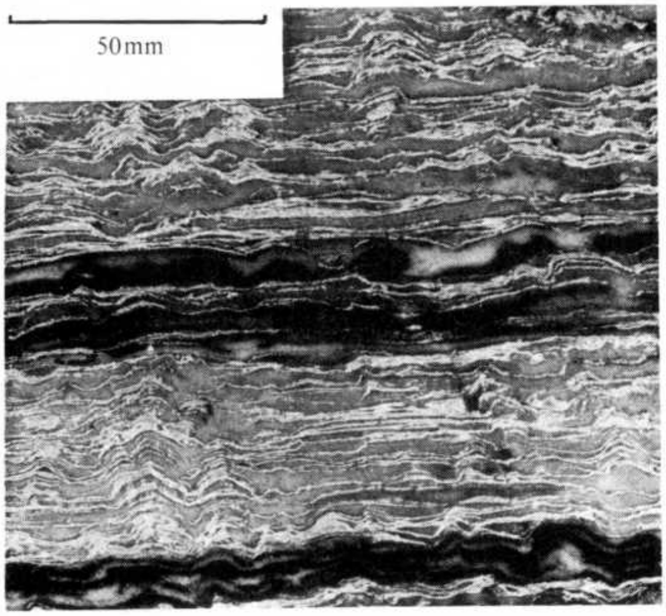


Fig. 3 Cut and etched surface of Mushandike stromatolitic limestone. Dark layers are of metasiltstone. Note preponderance of convex upward structures, moderate inheritance of domed laminae forming pseudocolumns in which successively younger laminae often develop a second order of curvature and thickening of laminae over the crests of flexures.

To the east of Gwenya hill, the sediments can be traced intermittently for some 40 km along the northern margin of the greenstone belt. For the most part the formation consists of ferrguinous, argillaceous sediments, banded iron-formation and a few small limestone pods. For much of this sector of the northern margin, the Gwenya formation is bounded below, and in places above, by the peridotitic eastern arm of the intrusive Mashaba Igneous Complex[6]. Approximately 20-km north-east of Fort Victoria, the Upper Greenstones and Mashaba Igneous Complex are cut by ~2,600 Myr Chilimanzi granite. On the Rockwood farm road, 4 km south-west of this contact, the Gwenya formation sediments are poorly exposed in unconformable contact with the underlying gneisses.

The older sedimentary rocks, of which the limestones are an integral infolded part, are an eastwards continuation of the Mashaba Sebakwian Group. Unlike the Gwenya formation rocks, which together with the overlying volcanic rocks of the Upper Greenstones have suffered only low-grade metamorphism, these sediments show widespread evidence of a medium grade, contact metamorphic event. Para-amphibolites, containing cordierite, show stellate clusters of anthophyllite, up to 100 mm across, cutting an earlier tectonic fabric. Relict, but occasionally fresh, andalusite has developed in metapelites adjacent to the limestones and in pelitic inclusions both within the Mushandike Granite and the surrounding gneisses[6]. When unaffected by the shearing (?~2,700 Myr), these andalusites have a random orientation. We attribute the development of these minerals to the thermal effects of the Mushandike Granite intrusion.

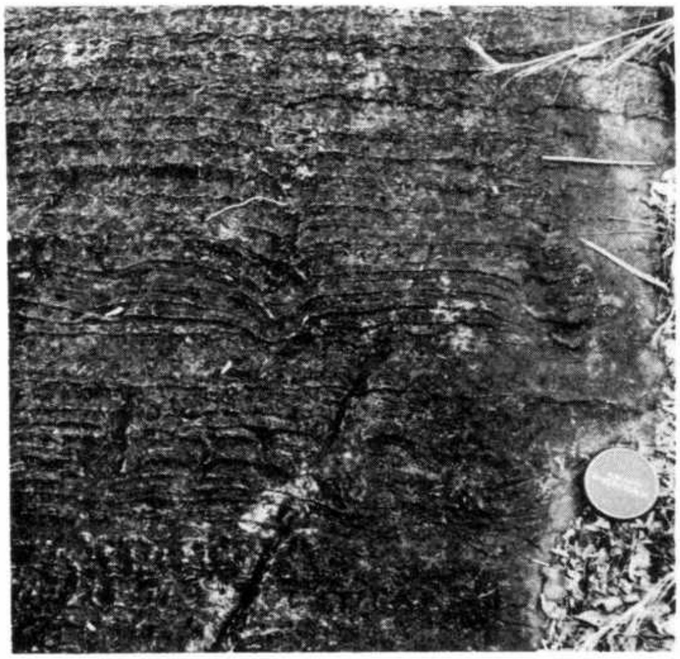


Fig. 4 Outcrop of Mushandike stromatolites showing large oblate domes with high laminae inheritance characteristics. Lens cap is 5 cm in diameter.

The limestone has a complex mineral assemblage, but commonly all specimens studied from widely spaced localities along strike contain the following: calcite, clinochlore, talc, tremolite, zoisite, grossularite and calcic plagioclase, which is compatible with the medium grade of the other Sebakwian rocks.

The stromatolites occur in a limestone unit that crops out over a 7-km strike, and has an average 200-m outcrop width (Fig. 1). Centimetre-thick, intercalations of metasiltstone are common. The limestone is in places massive, but is mostly finely laminated on a millimetre/centimetre scale, although the body has been tightly folded so that much of the lamination may be transposed foliation.

The stromatolite locality has escaped this intense deformation hence the biogenic structures have been fortuitously preserved.

The exposure is small and consists of a few scattered outcrops over an area of some 20 m^2. In these the millimetre-thick laminae are distinct, and following Hofmann[9] and Buick *et al.*[10], they are stratiform and corrugate, with a preponderance of convex-upwards structures in close to open spacing between laterally linked domes that have moderate inheritance. These form pseudocolumns, some of which develop a second order of curvature having the appearance of incipient dendroid branching (Fig. 3). Some of the larger domes are oblate and are formed from laminae with high inheritance characteristics (Fig. 4). In thin section the limestone has undergone extensive recrystallization, making recognition of microfossil structures hazardous. The morphology of the laminae, which characteristically thicken over the crests of flexures, does not resemble the plicated structure of stylolites, nor is it explicable as either a deformational or sedimentary structure.

The maximum height of any pseudocolumn is 3 cm, with the total relief of any lamina varying between 2 and 10 mm. The angle of inclination of a column axis to bedding is generally 90°, but may lean at 60°–70°. Hemispheroid radii are commonly up to 1 cm, but can be as much as 4 cm. The vertical increase in radius of successive hemispheroids is usually <0.1 mm, whereas the upward displacement of the centre of successive domes is 2 mm giving a growth factor of 0.2.

The relationship of the morphology described to named stromatolitic groups is uncertain, but the forms bear affinity to *Stratifera* and/or *Irregularia*.

The Mushandike stromatolite discovery is one of the most significant yet in the search for early life forms. To our knowledge they are the best example of ~3,500 Myr biogenic structures known and add to the growing body of evidence[11–14] that life was manifest as early as ~1,000 Myr after the formation of the Earth.

We thank A. Martin, M. J. Bickle and E. G. Nisbet for constructive criticism.

Received 12 February; accepted 8 April 1981.

1. Wilson, J. F. *et al. Nature* **271,** 23–27 (1978).
2. Wilson, J. F. *Spec. Publ. geol. Soc. S. Afr.* **5,** 1–23 (1979).
3. Wilson, J. F. in *The Precambrian of the Southern Hemisphere* (ed. Hunter, D. R.) (Elsevier, Amsterdam, 1980).
4. Martin, A., Nisbet, E. G. & Bickle, M. J. *Precamb. Res.* **13,** 337–362 (1980).
5. Bickle, M. J., Martin, A. & Nisbet, E. G. *Earth planet. Sci. Lett.* **27,** 155–162 (1975).
6. Wilson, J. F. *Rhod. geol. Surv. Bull.* **62,** (1968).
7. Hickman, M. H. *Nature* **251,** 295–296 (1974).
8. Coward, M. P. & Shackleton, R. M. *19th A. Rep. Res. Inst. Afr. Geol.* 55–58 (1975).
9. Hofmann, H. J. *Geol. Surv. Can.* Paper 69-39 (1969).
10. Buick, R., Dunlop, J. S. R. & Groves, D. I. *Alcheringa* (in the press).
11. Lowe, D. R. *Nature* **284,** 441–443 (1980).
12. Walter, M. R., Buick, R. & Dunlop, J. S. R. *Nature* **284,** 443–445 (1980).
13. Nisbet, E. G. *Nature* **284,** 395–396 (1980).
14. Nisbet, E. G. *Nature* **289,** 11–12 (1981).

Syn-sedimentary gravity slides (growth faults) in the Coal Measures of South Wales

T. Elliott & K. O. Ladipo

Department of Geology, University College of Swansea,
Singleton Park, Swansea SA2 8PP, UK

Growth faults are an important class of syn-sedimentary fault which have been widely recognized in sub-surface studies of hydrocarbon-bearing clastic basin successions[1–4]. They are listric normal faults which are active for a limited period during sediment deposition and therefore affect a discrete sedimentary interval. Whilst active, the faults substantially influence sedimentation and a thicker, generally sand-dominated succession accumulates on the downthrown side which is consistently towards the basin. The faults are not directly related to basement tectonics, and instead are thought to originate principally by gravity sliding within the sediment pile[5]. Overpressured/undercompacted conditions in buried muds aid development of the faults by reducing the slope required to initiate sliding[5], and in some cases associated mud diapirism also contributes by creating steeper slopes[6]. Despite their widespread recognition in sub-surface studies, relatively few examples have been described from exposed successions[7–10], and none from the economically important Upper Carboniferous Coal Measures of Europe. Here we describe examples of growth faults recently discovered in an opencast site of the Coal Measures in South Wales and comment on their possible significance.

The faults occur in the Middle Coal Measures (Westphalian B–C) which is the principal coal-bearing interval of the South Wales coalfield. They are exposed in the National Coal Board opencast site at Llanilid, north-west of Cardiff (NGR SS975813) in which the succession from the Five Foot-Gellideg Seam to the Bottom Abergorki Seam can be examined. This succession comprises 222 m of mudstones, siltstones and sandstones in which 14 workable coal seams occur. Growth faults have been detected at two levels in the upper part of this succession between the Gorllwyn and Bottom Abergorki Seams.

The lower set of faults affects an 8-m interval, and both underlying and overlying strata can be seen to be unaffected by the faulting (Fig. 1). The succession at this level comprises a series of small-scale coarsening-upwards sequences, two of which are affected by the faulting. The sequences commence with black, plant-rich shales which contain horizonated sideritic ironstone nodules. This facies coarsens upwards into siltstones with occasional thin (1–2 cm) sharp-based beds of fine-grained sandstone. The sandstones increase in thickness and frequency upwards and grade into a fine- to medium-grained sandstone unit. These uppermost sandstones contain abundant plant rootlets and ironstone nodules, and have a light-grey leached and mottled appearance typical of Upper Carboniferous sandy palaeosols. A coal seam overlies the fault-affected interval and is abruptly overlain by mudstones at the base of the next sequence. These coarsening-upward sequences reflect the repeated infilling of shallow-water environments, initially by the deposition of fine-grained sediment from suspension and subsequently by sand-laden currents. Rapid infilling of the area culminated in soil formation, plant colonization and peat/coal accumulation. The sequences may have accumulated in the interdistributary bays of a delta[11], or in shallow lakes or barrier-sheltered lagoons as small-scale deltas prograded into the area[12–15]. The lack of afficiated large-scale coarsening-upward sequences of delta front or beach-barrier origin at this stratigraphic level favours a lacustrine setting, perhaps analogous to the Atchafalaya area west of the present Mississippi delta[12].

Six faults are exposed within this interval in a horizontal distance of 165 m (Fig. 1). Each fault plane dips to the east and the strata are downthrown in this direction. The faults have a curved cross-sectional profile which is steep at the top (50°) and flattens progressively into a bedding plane fault in the shales at the base of the lower sequence and the underlying coal. The downthrown strata dip at angles up to 20° from the regional dip in the mid-depth region of the fault planes and terminate abruptly against the planes. The faults are thin zones only a few centimetres wide with minimal brecciation and only a weak foliation. In one example the distal portion of a fault steepens slightly from a bedding plane fault into a low angle thrust. Minor, westerly-dipping antithetic faults occur in the down-faulted mass and are typically steeper and straighter than the major faults. At the top of the faulted interval the dipping, faulted beds are truncated in the distal parts of each fault block and some of the material eroded from these sites appears to have been redeposited in depressions immediately adjacent to the fault on the downthrown side. The upper part of the fault-affected interval is overprinted by the leached palaeosol horizon at the top of the second sequence. These faults are clearly syn-sedimentary as both underlying and overlying strata are

0028–0836/81/210220—03$01.00

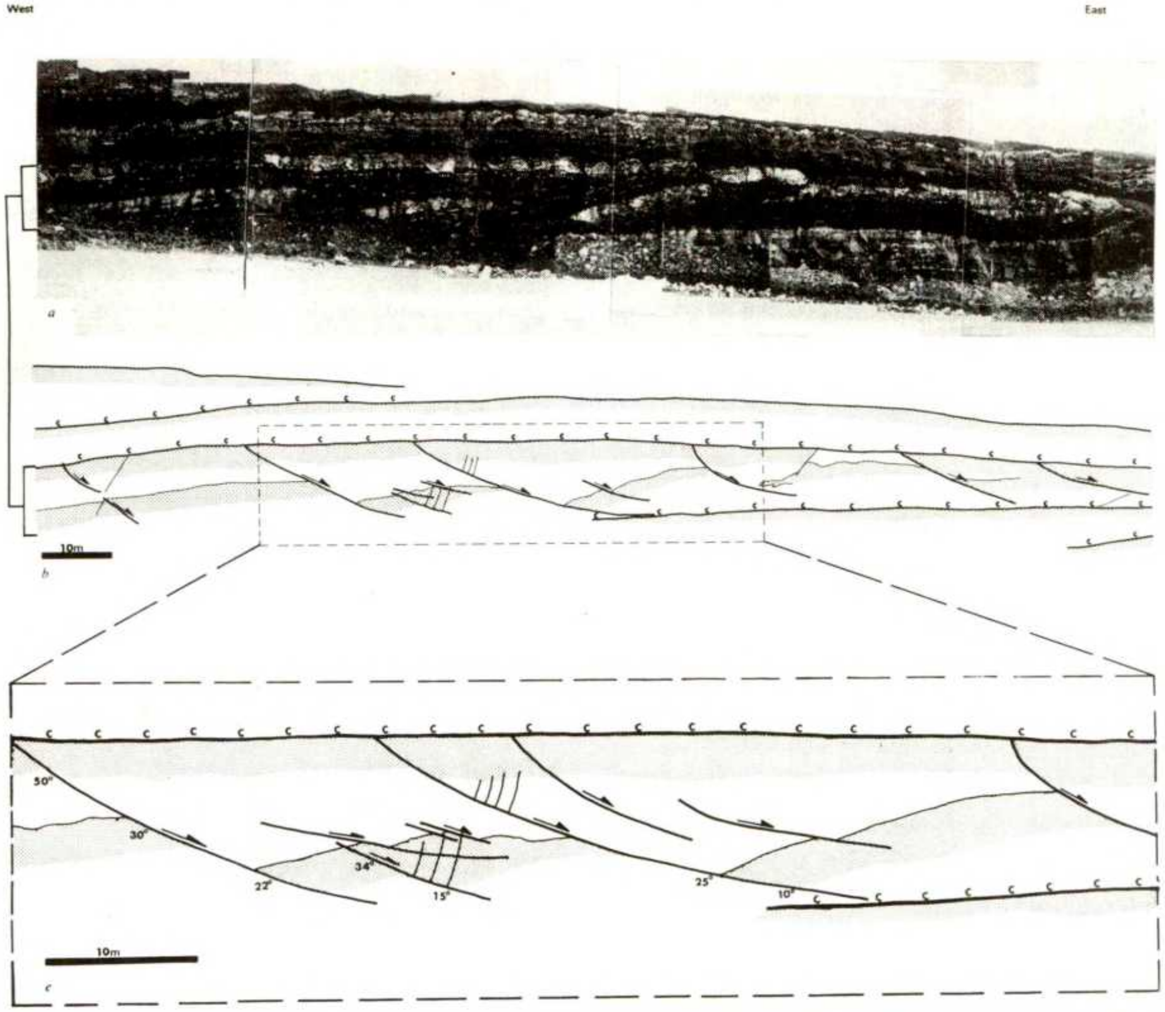


Fig. 1 *a*, Panorama of the lower set of growth faults in centre of photograph; *b*, line drawing of *a* showing the principal listric normal growth faults, occasional minor antithetic faults, the sandstones (stippled) and coal seams (c); note that both underlying and overlying strata are unaffected by the faults; *c*, detail of central part of *b* including dip angles of the faults.

unaffected, and the faulted strata were locally eroded before deposition of the next sequence. The major faults are listric normal faults reflecting easterly-directed gravity sliding. The orientation of the faults to the overall sediment transport direction is not known as the sedimentary structures have been obliterated by rootlets and palaeosol processes. The amount of deposition during faulting seems minimal, but is difficult to estimate in view of the small scale of the fault blocks and the extent of penecontemporaneous erosion in the distal parts of the fault blocks which has removed crucial evidence. However, it is clear that (? surface) relief created by the faulting was gradually obliterated by mass wasting processes possibly assisted by local wind-driven waves, resulting at least in re-sedimentation of this material. Following the planation of the area a period of soil formation and plant colonization ensued, overprinting the levelled surface of the fault blocks.

The upper set of faults affects the succession immediately below the Bottom Abergorki Seam. At this level the succession comprises thinly bedded, current ripple-laminated sandstones with interlaminated siltstones and mudstones which are cut by a series of lenticular, channelized, fine- to medium-grained sandstones. These facies possibly represent the upper parts of river mouth bar systems cut by the fluvial channels which supplied them[16], although research into this problem is still in progress. Three normal faults dipping generally eastwards are exposed in a horizontal distance of 45 m (Fig. 2). The principal fault has

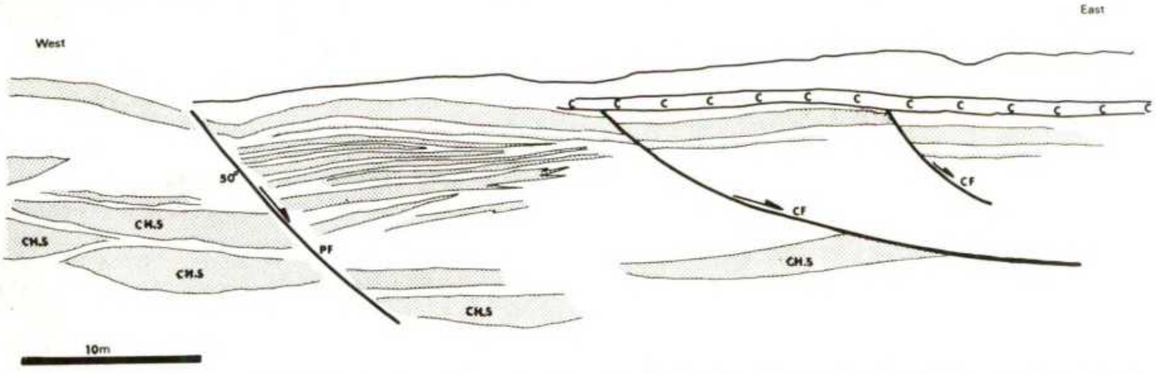


Fig. 2 The upper set of faults immediately below the Bottom Abergorki coal seam. The principal fault (PF) has only its upper part exposed, whereas the smaller scale, (?) crestal faults (CF) are completely exposed. Note the localized sandstone-dominated succession adjacent to the upper part of the principal fault which diminishes away from the fault. Sandstones are stippled. CH.S, channel sandstones; c, coal seam.

only its upper part exposed, dipping at 50° and flattening slightly with depth in the limits of exposure, suggesting a listric profile. This fault affects at least 18 m of succession and has a thicker, sandstone-dominated succession adjacent to the fault on the downthrown side. Numerous sandstone beds become thinner and finer grained when traced away from the fault plane, and eventually grade into siltstones and mudstones. These beds dip into the fault plane and are 'unconformably' overlain by the succeeding sandstone unit in the distal part of the fault block. Exposure of the fault plane permits the strike to be measured accurately at 072°/252°, indicating downthrow to the SSE. The preserved channel shape of numerous (inaccessible) sandstone units implies that the overall sediment transport direction was at a high angle to the strike of the faults. The localized sandstone-dominated facies on the downthrown side of the principal fault suggests that movements along this fault created a localized depression in the depositional area which became a prime site for the accumulation of sand-grade sediment. Differential sedimentation of this nature may have perpetuated the fault movements for some time causing the fault to grow until the locus of sedimentation shifted away from this area. On the downthrown side of the principal fault are two fully exposed, smaller scale faults with listric profiles. These faults parallel the principal fault and are downthrown slightly in the same direction, but do not exhibit a facies contrast between upthrown and downthrown sides. These smaller scale faults are clearly subordinate to the principal fault and are interpreted as minor, synthetic faults, possibly akin to the crestal faults of Evamy *et al.*[4] in the growth-faulted Tertiary depocentre of the Niger delta.

The described faults are therefore syn-sedimentary, listric normal faults which are interpreted as growth faults and are considered to have been produced by gravity sliding directed generally east or south-southeast. Evidence for sedimentation during faulting is limited for the lower small-scale faults, but is clearer for the upper, larger scale faults. Formation of the faults may have involved overpressured/undercompacted conditions in the mudstones which reduced the slope required to initiate sliding. Previously it was felt that relatively deep burial was required to produce overpressured conditions, but *in-situ* measurements of pore pressures in delta plain and delta front muds of the Mississippi delta have recently detected overpressured conditions at very shallow depths below the sediment surface[17,18]. In these conditions mass movement of sediment has occurred in the Mississippi delta on slopes as low as 0.14° (ref. 19), and Crans *et al.*[5] consider that the slope required to initiate gravity sliding diminishes to $<3°$. Slopes towards the lower end of this range of values may have existed on the small- to moderate-scale deltas in which the faults occur. As in the Mississippi delta, excess pore-water pressures may have been supplemented by methane gas (as a bubble phase?) in the richly organic muds of the Coal Measures.

The recognition of growth faults in the Coal Measures has considerable potential significance because if they prove to be common, particularly on a larger scale, they may influence the location, lateral extent, frequency, thickness and disposition of coal seams. Although the described faults are small by comparison with many of the sub-surface examples described from hydrocarbon provinces, they clearly influenced sedimentation and the resultant facies patterns (particularly the upper, larger scale example). The scale of the observed faults is in proportion with the scale of the exposures and larger scale examples may be recognizable in seam-correlated cross-sections through coalfields. A series of larger scale listric normal faults known collectively as the Jubilee Slide has been described in sub-surface studies of the South Wales coal basin in the Pontypridd area[20]. This structure comprises numerous slightly arcuate listric normal faults arranged *en echelon* in a fault zone traceable for 20 km along strike. The fault system affects the Lower, Middle and Upper Coal Measures and has a maximum throw of 90 m. The structure pre-dates later tectonic structures, but cannot be regarded as a syn-sedimentary growth fault as there is neither evidence that the faults influenced sedimentation, nor that they are overlain by beds not affected by the faulting. In view of the growth faults described herein the Jubilee Slide should be re-examined to determine whether or not it is a large-scale syn-sedimentary growth fault.

We thank the National Coal Board for access to opencast sites, and A. R. Gardiner and A. J. Pulham for comments on an earlier version of the manuscript. K.O.L. acknowledges a reasearch grant from the University of Ilorin, Nigeria.

Received 2 February; accepted 12 March 1981.

1. Ocamb, R. D. *Trans. Gulf-Cst Ass. Geol. Soc.* **11,** 139–175 (1961).
2. Busch, D. A. *Bull. Am. Ass. petrol. Geol.* **59,** 217–230 (1975).
3. Short, K. C. & Stauble, A. J. *Bull. Am. Ass. petrol. Geol.* **51,** 761–779 (1967).
4. Evamy, B. D. *et al. Bull. Am. Ass. petrol. Geol.* **62,** 1–39 (1978).
5. Crans, W., Mandl, G. & Haremboure, J. *J. petrol. Geol.* **2,** 265–307 (1980).
6. Bruce, C. H. *Bull. Am. Ass. petrol. Geol.* **57,** 878–886 (1973).
7. Edwards, M. B. *Bull. Am. Ass. petrol. Geol.* **60,** 341–355 (1976).
8. Chisholm, J. I. *Proc. Yorks. geol. Soc.* **41,** 305–323 (1977).
9. Rider, M. H. *Bull. Am. Ass. petrol. Geol.* **62,** 2191–2213 (1978).
10. Weimer, R. J. & Land, C. B. *Spec. Pap. geol. Soc. Can.* **13,** 633–666 (1976).
11. Elliott, T. *Sedimentology* **21,** 611–622 (1974).
12. Coleman, J. M. *Trans. Gulf-Cst Ass. geol. Soc.* **16,** 159–174 (1966).
13. Reading, H. G. *C.r. 6e. Congr. Int. Strat. Geol. Carbonif.*, Sheffield 1967 **IV,** 1401–1412 (1971).
14. Kanes, W. H. *Spec. Publ. Soc. econ. Paleont. Miner.* **15,** 78–106 (1970).
15. Donaldson, A. C. *Spec. Pap. geol. Soc. Am.* **148,** 47–78 (1974).
16. Wright, L. D. & Coleman, J. M. *J. Geol.* **82,** 751–778 (1974).
17. Bennett, R. H., Bryant, W. R., Dunlap, W. A. & Keller, G. H. *Mar. Geotech.* **1,** 327–335 (1976).
18. Hirst, T. J. & Richards, A. F. *Mar. Geotech.* **2,** 191–204 (1978).
19. Prior, D. B. & Coleman, J. M. *Mar. Geotech.* **3,** 37–60 (1978).
20. Woodland, A. W. & Evans, W. B. *Mem. geol. Surv.* **IV** (1964).

Visual observations of the Red Sea hot brines

A. S. Monin, E. A. Plakhin, A. M. Podrazhansky, A. M. Sagalevich & O. G. Sorokhtin

P. P. Shirshov Institute of Oceanology, USSR Academy of Sciences, 117218 Moscow, USSR

The research vessel Akademik Kurchatov has investigated the Red Sea rift zone with the aid of a manned submersible Pisces XI in the winter of 1979–80, as the value of such investigations is now well established[1,2]. Our research involved a comprehensive programme of hydrophysical studies in deeps containing thermal brines of the sea axial rift—Atlantis II, Discovery and Valdivia (see Fig. 1). Measurements were made with standard hydrographic stations, the AIST temperature and salinity meter with specially selected temperature and electric conductivity scales on board a towed vehicle Zvouk-4m, and Pisces XI. The present data on water stratification in the Atlantis II, Chain and Valdivia deeps have in general confirmed previously known data[3–9] but have also revealed some new facts. For example, we have recorded an increase of temperature of the hot brine of up to 62 °C in the southwestern basin of the Atlantis II deep since studies aboard the R/V Sonne in 1977. Using data from this expedition, Hartmann[9] gives a curve of temperature growth in the lower brine constructed from the data of different expeditions since 1965. A multilayer temperature structure of the upper brine and horizontal inhomogeneities have also been found[8–10]. Steps in the upper brine of the Atlantis II deep were first described [6,9] from the surveys of the R/V Valdivia (1971–72) and Sonne (1977). Our observations reveal stratification of the deep brine (Table 1) in temperature and electrical conductivity (salinity) and enable their rate of change of thickness to be estimated[10].

Water in the area of the deeps and on the relief saddle between the Chain and the Discovery deeps was sampled using a 15-l water bottle designed to function at 5–10 m depths above the sea floor. The samples were analysed hydrochemically (Table 2) as well as in a spectrohydronephelometer which measured spectral transparency curves.

A comparison of these curves for the bottom water shows that in the Atlantis II deep a sharp decrease in transparency is

0028-0836/81/210222—04$01.00

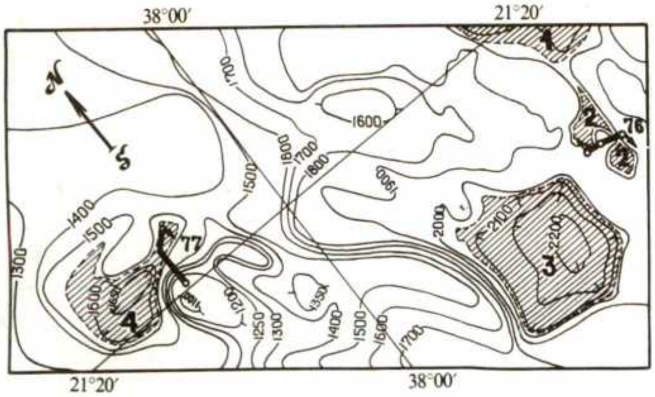


Fig. 1 A scheme of the "Pisces XI" dives into the Chain (76) and Valdivia (77) deeps: 1, Atlantis II deep; 2, Chain deep; 3, Discovery deep; 4, Valdivia deep.

observed at ~11 m above the bottom (for $\lambda_{ef} = 546$ nm): a 10-m thick turbid bottom layer seems to exist. The transparency curve of the Valdivia deep resembles the curve from the Atlantis II deep; however, in the Discovery deep the bottom water differs appreciably. On the saddle between the Chain and the Discovery deeps the optical properties of the water are intermediate between those of the Valdivia and the Discovery deeps. Note that the Discovery and the Atlantis II waters differ considerably in their chemical composition (Table 2). Our analyses indicate that the nature of the brines changes within the deeps.

Taking into consideration the large amount of data available for this region, our expedition aimed to verify the hypothesis of a common source of thermal water in the Atlantis II, Chain and Discovery deeps[4,7,8] and to investigate the Valdivia deep for which only scanty data have been published[5].

The hypothesis of the formation of high-salinity thermal water in the Chain and the Discovery deeps assumes an overflow of this water through bottom depressions from the southwestern basin of the Atlantis II deep.

Table 1 Thermohaline characteristics of water in brine steps within the Atlantis II deep

No. of steps	Thickness (m)	Temperature (°C)	Salinity (%)
1	4	40.9	67.6
2	8	47.1	113.6
3	16	51.5	144.3
4	10	52.2	145.1
5	From 2,049m depth to the bottom	62.3	275.0

The Pisces XI submersible was equipped with hydrophysical sensors and a dive was made to the supposed place of contact of thermal waters between the Chain and the Discovery deeps (see Fig. 1, route 76). Meridional towing of the Zvouk-4m vehicle across the saddle had already yielded 14 photographs of the saddle bottom (see, for example, Fig. 2). The photographs show a highly dissected topography with rock outcrops and numerous fragments of apparently basaltic rocks powdered with sediment often white and showing ripple marks in some areas. Figure 2 *a*, *b* indicates a complicated bottom topography within the saddle. Steep step-like scarps of a 3–4 m amplitude have been fixed (Fig. 2*a*). In the vertical walls, bedrock crops out, flat areas are powdered with loose sedimentary material. Judging by the bedrock fragments, we have observed, not effusive formations encountered in the axial rift zone, but carbonate strata (marls) or iron–manganese consolidated sediments the fragments of which were found in the submersible's ski (dive 76). A tectonic origin of the scarps is most probable. Ripples are seen on the scarp to the right of the rock in Fig. 2*a*, on the left-hand part of the photograph in Fig. 2*b* and on the whole of Fig. 2*c* which reveals mysterious decimetre-scale ripples (possibly, on the brine surface).

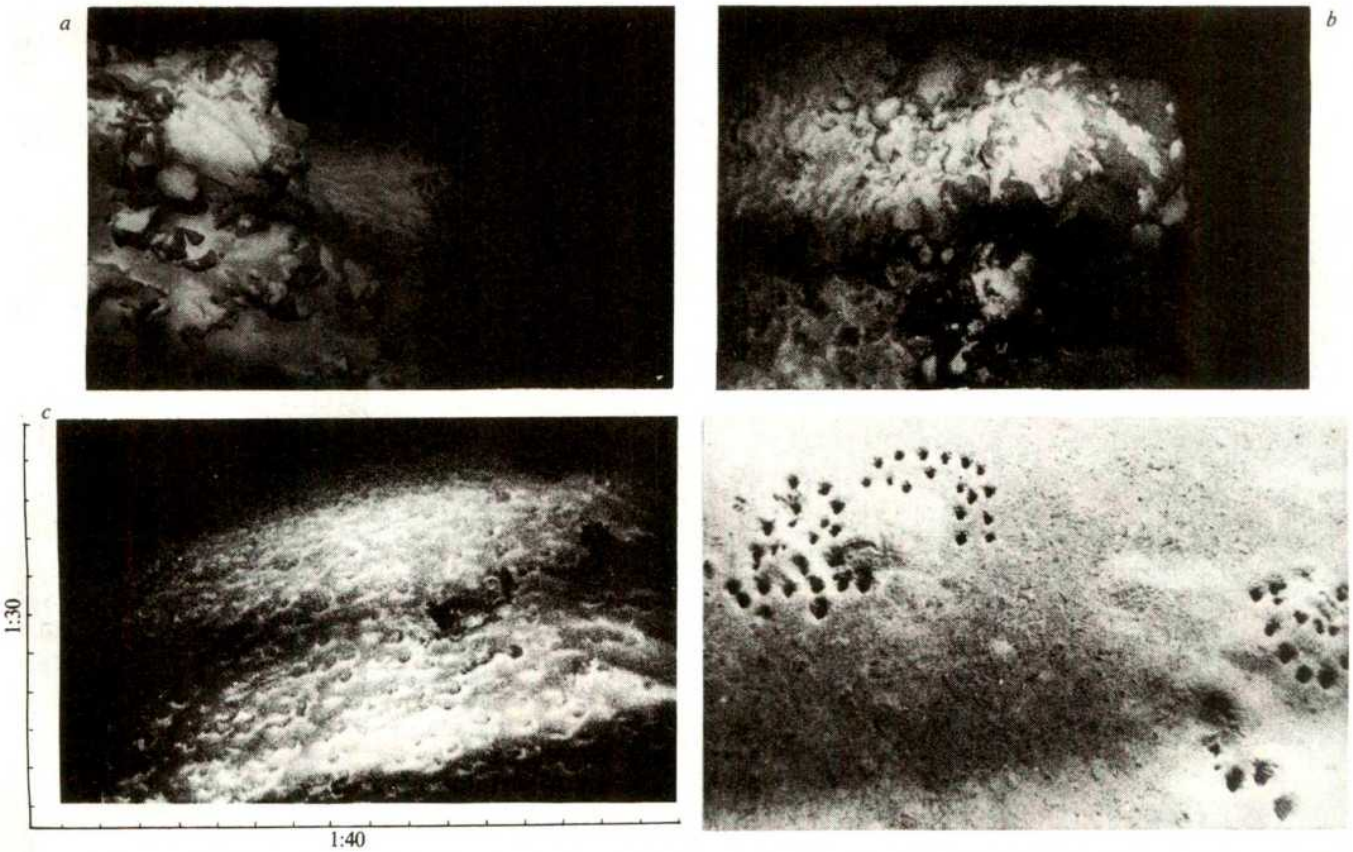


Fig. 2 Bottom and brine surface photographs made in the diving area with the aid of the towed vehicle Zvouk-4m: *a*, bottom area covered with ripples (to the right of the rock); *b*, bottom area with ripples (in the left-hand part of the photograph); *c*, decimetre-scale ripples (possibly, on the brine surface); *d*, a group of bottom holes.

Table 2 Hydrochemical characteristics of deep brines and surface water in the Red Sea

	Brine level (m)	Transitional layer (m)	Sampling depth (m)	Temperature (°C)	Salinity (‰, conductometric, at bulk dilution)	Oxygen (O_2 ml per l)	Si (μg atom l^{-1})	Disposable CO_2 (ml per l)	Total CO_2 (ml per l H_2O)
Atlantis II									
Sample 1	2,028	1,975	2,127	62	321	0	800	5.6	44
Sample 2	2,056	2,052	2,070	62	322	0	—	5.8	43
Discovery	2,060	1,995	2,118	44	315	0.2	150	0.21	9.2
Valdivia	1,565	1,560	1,657	32	286	0	140	2.5	83
Red Sea	—	—	0	27	38.9	4.5	11	0.21	51

The Pisces XI dive (76) onto the Chain–Discovery saddle on 12 February 1980 showed that the water changed its transparent-blue colour to dull yellow (due to a high concentration of suspended matter in the upper part of the sharp pycnocline layer) at a depth of 1,995 m, as the bottom was approached. The water temperature increased from 21 to 25 °C and the sound velocity from 1,560 to 1,572 m s^{-1}. The sea floor at a depth of 2,005 m was covered with light-yellow sediment with easily turbidable black coating, straight parallel sand bars 1.5–2 m apart and with separate white spots, decimetres in diameter, and centimetre-high knolls (possibly, evaporite outcrops). Although there were no traces of life on the bottom a dead fish found there may be indicative of the conditions preserving organic matter, that is the absence of oxygen and bacteria.

The submersible moved towards the Discovery deep, just above the apparent bottom which looked to the observers like mud covered with waves. However, in their attempts at touching-down on this bottom the observers made sure that they saw (and the echosounder of the Pisces recorded) the brine surface.

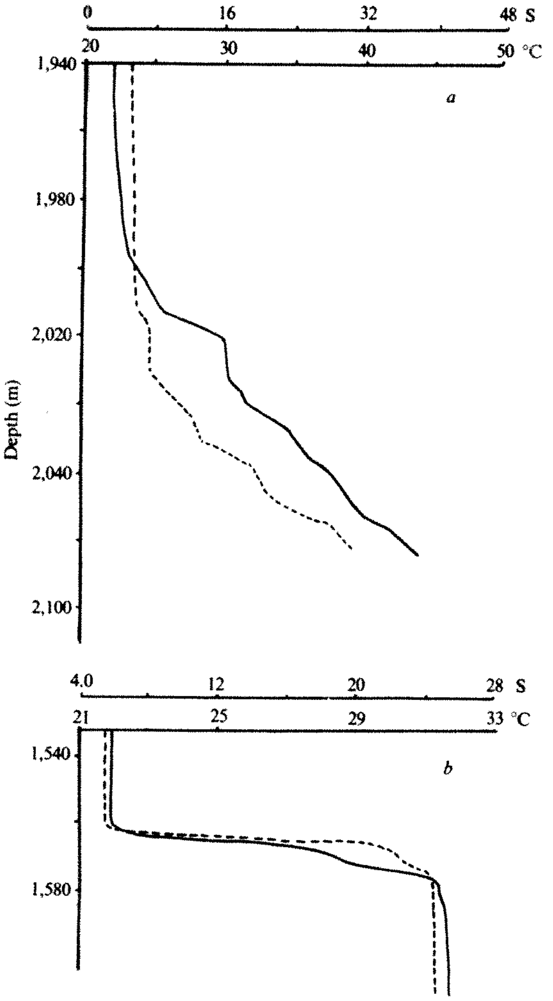


Fig. 3 Curves showing vertical distribution of temperature (solid line) and conductivity (dashed line) in the thermal brine layer of the Chain (*a*) and Valdivia (*b*) deeps.

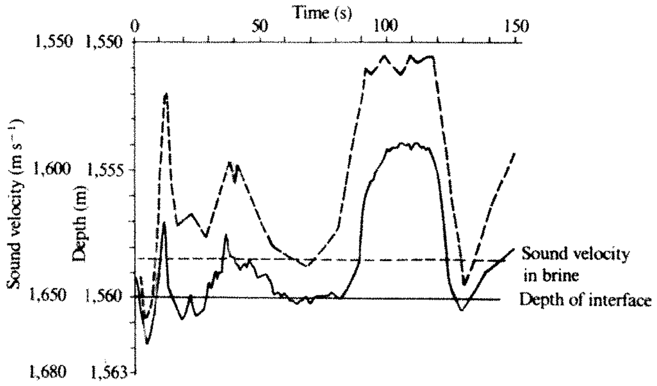


Fig. 4 Graphic recordings of sound velocity and depth during the Pisces XI manoeuvres in its dives into the Valdivia deep (a sample of the recordings of the submersible's sensors during 150 s).

Going deeper into the brine the water became sharply turbid, a streaming yellowish haze could be seen through the portholes, overboard temperature increased to 33 °C and sound velocity to 1,625 m s^{-1} (the Chain deep). At the maximum depth of 2,030 m the buoyancy force of the pycnocline balanced the forces of the thrusters and stopped the descent of the submersible.

The analysis of the hydrophysical and echosounder recordings made during the crossing of the saddle shows no evidence of brine overflows from the Chain to the Discovery deeps. Contact between the deeps takes place, evidently, only at the depth of the transition layer between the normal Red Sea deep waters and the brine layer (1,980–2,010 m layer, Fig. 3*a*).

Hydrophysical studies of the Valdivia deep (Fig. 1, route 77) were also undertaken. Preliminary standard hydrographic casts and soundings with the AIST had revealed a sharp pycnocline with a very thin transition layer (Fig. 3*b*). During the dive into the Valdivia deep, on 14 February, 1980 the submersible reached the bottom at a depth of brine of 1,530 m east of the deep. On setting course for the west, at a depth of 1,555 m the Pisces entered the transition layer which was immediately noted by the observers as a streaming viscous haze. A pycnocline layer as a turbid surface covered with ripples was clearly seen 3–5 m below the submersible. As in the dive to the saddle described above, no penetration through the pycnocline was possible, and after attempting to go deeper with the aid of its thrusters the submersible headed eastwards above the brine surface. The manoeuvres of the submersible during its attempts to go deeper into the brine are shown in Fig. 4 together with the results of continuous sound velocity measurements in the brine (a sample taken over 2.5 min is presented).

Moving above the brine surface the submersible reached its underwater 'shore' covered with yellowish mud and descended at 30–40° into the underwater salt 'lake'. Against the background of the slope the brine level was distinct, resembling a whitish mist surface, and the observers noted underwater waves at a depth of 1,560 m (internal waves on the brine surface which was disturbed by the Pisces thrusters) which were rolling on the shore of the salt lake and then were rolling down in streams.

Furthermore the Pisces started ascending along the bottom rise slope covered with sediments with light-yellow and brown crests of thin-layered sedimentary rocks usually inclined towards the slope and covered with carbonate crusts with a dark yet thin manganese 'tarnish' on the surface. The inclined bedding of sedimentary rocks is most likely to be due to slide phenomena on the salt-bearing slopes of the Red Sea main trough.

According to M. S. Barash (personal communication) the samples taken during the dives from laminated sedimentary rocks contain abundant well-preserved late Quaternary planktonic foraminifers (*Orbulina universa* d'Orb; *Globigerinita glutinata* Egger; *Globigerinoides sacculifer* Brady; *Globigerinoides ruber* d'Orb; *Globigerinella aquilateralis* Brady; *Globigerina rubescens* Hofker; *Globigerinoides tenellus* Parker).

Ushakova studied the same samples and determined Pleistocene (*Emiliania huxleyi* Lohman; *Unibilicosphaera mirabilis* Lohman; *Geophyroacapsa oceanica* Kamptner; *Coccolithus atlanticus* Cohen) redeposited Cretaceous (*Lucianorh abdus cayeuxi* Deflandre; *Microrhabdulus decoratus* Deflandre; *Lithraphidites quadratus* Bramlette, Riedel) and Palaeogene calcareous nannoplankton (*Helicopontosphaera intermedia* Martini, Hay, Mohler; *Discoaster multiradiatus* Bramlette, Riedel; *Pseudozygrhablithus latus* Hag; *Marcalius inversus* Deflandre, Bramlette, Martini).

The carbonate crusts covering thin-laminated sedimentary rocks are identical in composition to those studied in detail in the Red Sea rift zone at 18 °N where their formation is associated with the climate cooling stages in Pleistocene. These crusts are most likely to be dated as Würmian. Their origin could be accounted for by a partial bridging of the Bab-el-Mandeb strait due to a general lowering of the World Ocean level during Würmian glaciation. Several metres above the lake level, separate pink shrimps and numerous groups of holes (1–3 cm in diameter and ~1 dm apart) were noted in the sedimentary cover, the latter being found almost everywhere on the Red Sea bottom (see, for example, Fig. 2*d*). No such groups of holes were encountered in the vicinity of the brine surface or on the bottoms of the thermal deeps. This may be evidence for their biological origin.

Received 12 September 1980; accepted 18 February 1981.

1. Riffaud, C. & Pichon, X.Le, *Expédition "Famous"* (Albina Michel, Paris, 1976).
2. CIAMEX, *EOS* **59,** 1198 (1978).
3. Voitov, V. I. *Dokl. Akad. Nauk SSSR,* **174,** 1429–1431 (1967).
4. Bubnov, V. A. *et al. Okeanologia,* **17,** 603–611 (1977).
5. Backer, H. & Schoell, M. *Nature phys. Sci.* **240,** 153–158 (1972).
6. Schoell, M. in *Proc. 2nd United Symp. on Development and Use of Geothermal Resources,* San Francisco (1975).
7. Degens, E. & Ross, D. *Hot Brines and Recent Heavy Metal Deposits in the Red Sea.* (Springer, Berlin, 1969).
8. Booker, H. *Geol. Jb.* 3–78 (1975).
9. Hartmann, M. *Deep-Sea Res.* **27** A, 161–171 (1980).
10. Monin, A. S., Plakhin, E. A. & Prokhorov, V. I. *Dokl. Akad. Nauk. SSSR* **254,** 483–486 (1980).
11. Monin, A. S. *et al. Dokl. Akad. Nauk SSSR* **254,** 1005–1009 (1980).
12. Monin, A. S., Plakhin, E. A. & Stunjas, P. A. *Dokl. Akad. Nauk. SSSR* **255,** 458–462 (1980).

Atlantic-type carbonate stratigraphy in the late Miocene Pacific

Dean A. Dunn, T. C. Moore Jr
& Lloyd D. Keigwin Jr*

Graduate School of Oceanography, University of Rhode Island,
Kingston, Rhode Island 02881, USA

In previous attempts to correlate fluctuations of climate with the carbonate content of deep-sea sediments, using the ratio ^{18}O/^{16}O as an indicator of palaeotemperature, conflicting patterns have emerged between Pacific and Atlantic deep-sea cores. Although there is controversy about the relative importance of the primary productivity of calcareous microorganisms[1-3] and their dissolution at depth[4-10], in Pacific Pleistocene cores high-carbonate sediments correlate with glacial episodes and low-carbonate sediments with interglacials[1-4,7-10]. In Atlantic Pleistocene cores, on the other hand, the correlation seems to be reversed[4,5,7,10-12]. We report here that a late Miocene (6.14–6.53 Myr) deep-sea sediment core from the eastern equatorial Pacific (DSDP Site 158) shows a negative correlation between carbonate content and benthic oxygen isotopes which is similar to that found in Quaternary records of the Atlantic rather than the Pacific. The explanation may be that Atlantic and Pacific waters differed in HCO$_3^-$ concentration at depth as they do today, and that in the late Miocene, Atlantic water flowed over the submarine sill at the Isthmus of Panama and at least periodically, reached the location of Site 158.

Stable isotope and carbonate analyses were performed on 44 sediment samples from DSDP Site 158 (Cocos Ridge, eastern equatorial Pacific Ocean, 06°37.36′ N, 85°14.16′ W, 1,953 m water depth). Samples were taken from cores 15 and 16 at intervals of 20 cm across the late Miocene ^{13}C depletion in benthic foraminifera[13,14], and thus, our samples span an interval of 390 kyr between ~6.14 and 6.53 Myr, based on the estimated age of 6.3 Myr for the carbon shift[15]. The sample spacing represents an interval of ~10 kyr between samples, based on the estimated sedimentation rate of 2.0 cm kyr^{-1} for the late Miocene[16]. Stable carbon and oxygen isotopic ratios of the benthic foraminifera[14] *Globocassidulina subglobosa*, and benthic/planktonic (B/P) foraminiferal ratios were determined for each sample. Carbonate percentages for each sample were determined by the revised carbonate bomb techniques[17].

Our results for Site 158 show that the highest percentages of carbonate correlate with lighter benthic oxygen isotopic ratios (Fig. 1*a*), indicating warmer palaeotemperatures and/or less continental ice volume. The maximum correlation occurs with an offset of one sample interval, suggesting that changes in the carbonate content lead changes in the oxygen isotopic content by ~10 kyr. This conflicts with previous studies of Pleistocene equatorial Pacific cores[18-21] which estimated that carbonate oscillations lagged behind oxygen isotope variations by 5–10 kyr. The pattern of Site 158 is that of an Atlantic-type carbonate stratigraphy (high-carbonate interglacials); although the maximum magnitude of the correlation coefficient (0.42 at one sample lag) is low, it is of approximately equal magnitude with the maximum oxygen isotopic–carbonate correlations of Pleistocene cores V19–29 ($r = 0.52$) and V28–238 ($r = 0.35$) from the equatorial Pacific[21], despite the fact that cores V19–29 and V28–238 show the opposite Pacific-type pattern.

Our data also show that higher B/P values correlate with lower carbonate percentages (Fig. 1*b*). The maximum correlation coefficient ($r = 0.54$) occurs at one sample lag, indicating that changes in carbonate content lead corresponding changes in B/P ratio by ~10 kyr, and that B/P ratio and oxygen isotope records are in phase. Because dissolution preferentially removes planktic foraminifera, higher B/P ratios indicate increased dissolution. The lag between these two indicators of dissolution is perplexing. Total carbonate content (both coccoliths and foraminifera) may respond more rapidly to changes in bottom water chemistry than do changes in foraminiferal faunas alone. The correlation of B/P ratios with per cent carbonate suggests that in the late Miocene (6.14–6.53 Myr) samples from DSDP Site 158, carbonate percentage is controlled primarily by dissolution. Furthermore, it indicates that times of greater δ^{18}O ('glacials') are times of greater carbonate dissolution, the pattern found in Pleistocene Atlantic deep-sea cores[4,5,7,10-12].

Spectral analysis of such short time series, especially those for which the time-spacing of samples is only approximate, cannot

---

*Present address: Department of Geology and Geophysics, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA.

0028-0836/81/210225—03$01.00

be expected to give very accurate estimates of the variance distribution as a function of frequency; however, such an analysis can give a general idea of how the variance is partitioned between the low and high frequency parts of the spectrum. More important, cross-spectral analysis[21,22] of the carbonate and oxygen isotope paired time-series can provide a useful indication of the frequencies over which the two records are most highly correlated. The age model for our spectral analyses is based on the shift in benthic ^{13}C isotopes at 136.50-m depth in this site, which has been dated[15] at 6.3 Myr, and the estimated sedimentation rate[16] of 2.0 cm kyr^{-1}.

The oxygen isotope spectrum shows broad spectral peaks centred at frequencies corresponding to periods of 167 kyr, 53 kyr, and 22.7 kyr (Fig. 2). The carbonate spectrum shows broad peaks at 235 kyr and 29.4 kyr, with only a shoulder in the spectrum near the frequency corresponding to a period of ~60 kyr. These results agree with a study of the carbonate record[23] from the entire late Miocene of Site 158 which showed that most of the variance in the record is associated with periods of oscillation longer than 150 kyr. The presence of spectral peaks in both records near the periods associated with the changes in obliquity and precession of the Earth's orbit[24] may be only fortuitous.

A high degree of correlation (coherency) occurs between the oxygen isotopic and the carbonate signals of Site 158 at low frequencies (periods longer than 111 kyr, with a coherency peak at a period of 182 kyr). A second region of high coherency is centred about a period of 30 kyr. The coherence between the two curves at long periods of oscillation may be due to matching a few peaks and valleys in a relatively short section of the carbonate record. However, the length of record (~390 kyr) is comparable with many of those which have been used to establish the relationship between carbonate and isotope stratigraphies in the Pleistocene[18–21]. Moreover, the coherency between the oxygen isotopic record and the carbonate record even at higher frequencies lends credence to the proposed relationship between the two records.

The discovery of an 'Atlantic-type' carbonate stratigraphy in the late Miocene eastern equatorial Pacific Ocean and the eventual development of the 'Pacific-type' stratigraphy may be related to the tectonic evolution of the Isthmus of Panama. The carbonate variations seen in the late Miocene at Site 158 are primarily due to dissolution, as shown by the B/P foraminifera ratios, with less dissolution occurring during the warmer (isotopically lighter) intervals. This pattern may have been caused by the exchange of water between the Atlantic and Pacific through the open seaway of Panama. Thus, before the middle Pliocene closure of the Isthmus of Panama[25,26], periods of warming (lower $\delta^{18}O$) may have been associated with increased flow of Atlantic waters across the Panama Seaway and into the Pacific basin. This water would be less corrosive to calcite because it would be younger and contain less dissolved CO_2 than

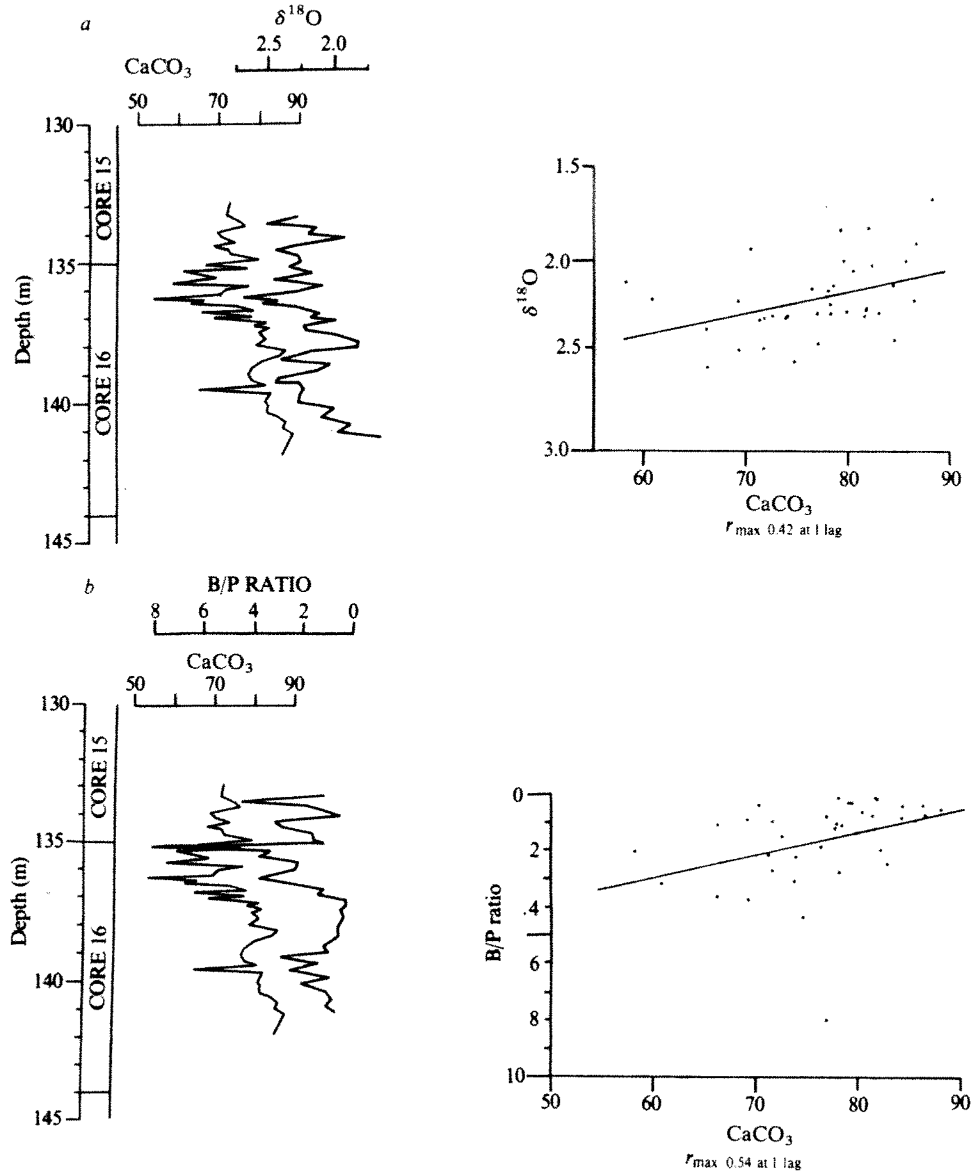


Fig. 1 *a*, Carbonate percentages and stable oxygen isotopic ratios of the benthic foraminiferan *Globocassidulina subglobosa* (relative to the B-1 powder standard) for 44 paired samples in DSDP Site 158. Per cent carbonate and $\delta^{18}O$ are negatively correlated, implying that there are high-carbonate warm intervals (an 'Atlantic-type' carbonate stratigraphy) in this sediment from the late Miocene eastern equatorial Pacific. The *y*-axis ($\delta^{18}O$) has been inverted in this figure to make the Atlantic-type correlation more obvious. The maximum correlation coefficient ($r = 0.42$) occurs at one sample lag, indicating that changes in benthic oxygen isotopic content lag behind changes in carbonate percentage by ~10 kyr. *b*, Carbonate percentages and B/P foraminiferal ratios for DSDP Site 158 samples. Per cent carbonate and B/P ratios are negatively correlated, implying that variations in carbonate percentage are primarily due to dissolution, as shown by high B/P ratios correlating to low carbonate percentages. Maximum correlation coefficient ($r = 0.54$) occurs at one sample lag, indicating that changes in the B/P ratio lag behind corresponding changes in carbonate percentage by ~10 kyr.

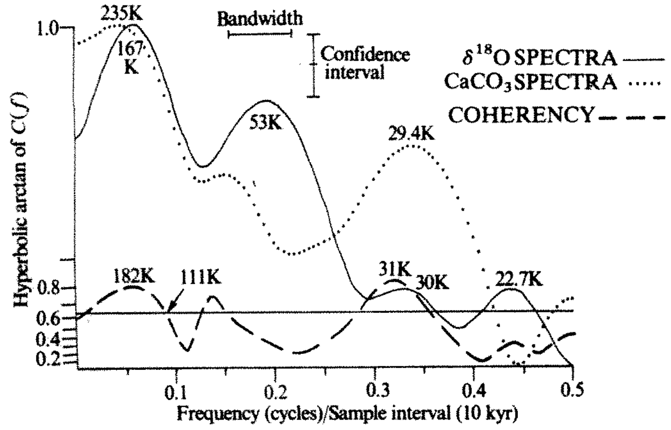


Fig. 2 Spectral analysis of stable oxygen isotopic ratios of *Globocassidulina subglobosa* and carbonate percentage, and coherency between these two signals. Horizontal axis is cycles per 10 kyr. Log spectra of both time series are plotted on the vertical axis as power (equal to mean variance squared) versus frequency. Spectral estimates for δ^{18}O and CaCO$_3$ have been scaled by a constant, so that the area under each curve equals one. Vertical axis for coherency (hyperbolic arctangent of coherency for each frequency) was chosen so that the confidence interval is constant for all frequency bands.

the Pacific waters. During cool intervals, Pacific water masses (which have been away from the surface longer and thus are higher in dissolved CO$_2$ and more corrosive to calcite) could have replaced those derived from the Atlantic. The exact mechanism which might drive this oscillation in the source of deep waters for the eastern tropical Pacific is uncertain, but it may be associated with differing rates of production of deep waters in the Northern and Southern Hemispheres.

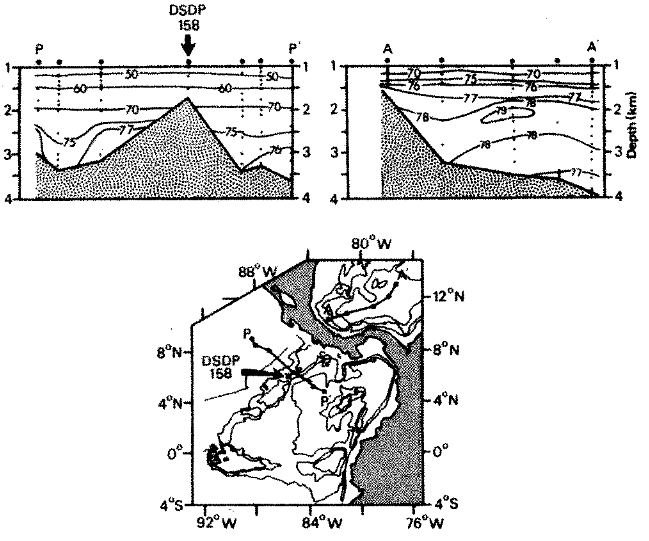


Fig. 3 Modern hydrographical sections for the Colombian Basin and the eastern equatorial Pacific. Density contours are plotted as values of $(\sigma_\theta - 27) \times 100$; a σ_θ value of 27.70 would be represented by 70. The estimated late Miocene palaeodepth of Site 158 is 1,700–1,800 m (ref. 31). Assuming that late Miocene density stratification of these basins was similar to the modern situation, then water masses passing over a minimum sill depth of 1,000–1,200 m would sink to the depth of Site 158.

At present, surface waters on opposite sides of the Panama Isthmus have similar surface temperatures (26–28 °C), but Atlantic waters have higher salinities. Flow of warm, saline Caribbean water from the Atlantic North Equatorial Current into the eastern equatorial Pacific has been proposed as a contributing factor to accumulation of Pliocene carbonate-rich sediments in the equatorial Pacific[27]. Also, passage of shallow Caribbean waters across Panama in the Pliocene has been implied as the mechanism for migration of a shallow-water benthic foraminiferan (*Amphistegina gibbosa*), normally restricted to the Atlantic–Caribbean province, into sediments off the California Continental Borderland[28]. Examination of modern hydrographic data[29,30] from the western Colombian Basin and the eastern equatorial Pacific Ocean (Fig. 3) indicates that Atlantic water masses passing across the Panama sill would sink to greater depths in the eastern Pacific, due to high densities found at shallower depths in the Atlantic. (Additional hydrographic data for the Colombian Basin and the eastern equatorial Pacific was made available by the National Oceanographic Data Center.) Thus, flow of Atlantic waters westward into the Pacific is theoretically possible, with circulation being driven by sinking of Atlantic water masses along isopycnal surfaces to greater depths in the Pacific.

Based on an estimated back-tracked depth of 1,700–1,800 m for Site 158 during the late Miocene[31], and assuming that density stratification of late Miocene ocean basins was similar to the modern situation, we estimate that the minimum depth for the late Miocene Panama sill was 1,000–1,200 m. This sill depth would allow Atlantic intermediate water masses to be transported through the Central American seaway. In the eastern Pacific, the Atlantic water masses could sink to the estimated depth of Site 158, causing the reduced carbonate dissolution at this site and producing the 'Atlantic-type' carbonate stratigraphy. If this proposed hypothesis is valid, then the shoaling of the Panama Seaway should mark the end of the negative correlation between the oxygen isotope and carbonate records.

We thank J. P. Kennett, N. G. Pisias, and W. L. Prell for critical reviews, K. Elmstrom and S. Hagerty for laboratory assistance, and M. B. Carter and N. G. Pisias for computer programming advice. Research was supported by NSF through grants to CENOP (OCE 76-81489 and OCE 79-14594).

Received 17 November 1980; accepted 26 February 1981.

1. Arrhenius, G. *Swedish Deep-Sea Expeditions (1947–1948)* Rep. **5**, 227 (1952).
2. Valencia, M. *Bull. geol. Soc. Am.* **88**, 143–150 (1977).
3. Adelseck, C. G. Jr & Anderson, T. F. *Geology* **6**, 388–391 (1978).
4. Berger, W. H. *J. Foram. Res.* **3**, 187–195 (1973).
5. Olausson, E. *Prog. Oceanogr.* **3**, 221–252 (1965).
6. Berger, W. H. *Deep-Sea Res.* **15**, 31–43 (1968).
7. Broecker, W. S. in *The Late Cenozoic Glacial Ages* (ed. Turekian, K. K.) 239–265 (Yale University Press, New Haven, 1971).
8. Hays, J. D., Saito, T., Opdyke, N. D. & Burckle, L. H. *Bull. geol. Soc. Am.* **80**, 1481–1514 (1969).
9. Thompson, P. R. & Saito, T. *Geology* **2**, 333–335 (1974).
10. Gardner, J. V. *Cushman Fdn Foram Res Spec. Publ.* **13**, 129–141 (1975).
11. Olausson, E. in *The Micropalaeontology of Oceans* (eds Funnell, B. M. & Riedel, W. R.) 375–398 (Cambridge University Press, 1971).
12. Olausson, E. *Prog. Oceanogr.* **4**, 245–265 (1967).
13. Keigwin, L. D. Jr *et al.* (in preparation).
14. Keigwin, L. D. Jr *Earth planet. Sci. Lett.* **45**, 361–382 (1979).
15. Keigwin, L. D. Jr & Shackleton, N. J. *Nature* **284**, 613–614 (1980).
16. van Andel, T. H. *et al. Init. Rep. DSDP* Leg **16** (1973).
17. Dunn, D. A. *J. Sed. Petrol.* **50**, 631–637 (1980).
18. Luz, B. & Shackleton, N. J. *Cushman Fdn Foram. Res. Spec. Publ.* **13**, 142–150 (1975).
19. Ninkovich, D. & Shackleton, N. J. *Earth planet. Sci. Lett.* **27**, 20–34 (1975).
20. Shackleton, N. J. & Opdyke, N. D. *Geol. Soc. Am. Mem.* **145**, 449–464 (1976).
21. Moore, T. C. Jr, Pisias, N. G. & Heath, G. R. in *The Fate of Fossil Fuel CO$_2$ in the Oceans* (eds Andersen, N. R. & Malahoff, A.) 145–165 (Plenum, New York, 1977).
22. Pisias, N. G., Dauphin, J. P. & Sancetta, C. *Quat. Res.* **3**, 3–9 (1973).
23. Dunn, D. A. & Moore, T. C. Jr *Bull. Geol. Soc. Am.* **92**, Part II, 408–451 (1981).
24. Berger, A. L. *Nature* **269**, 44–45 (1977).
25. Saito, T. *Geology* **4**, 305–309 (1976).
26. Keigwin, L. D. Jr *Geology* **6**, 630–634 (1978).
27. Emiliani, C., Gartner, S. & Lidz, B. *Palaeogeog. Palaeoclimatol. Palaeoecol.* **11**, 1–10 (1972).
28. Crouch, R. W. & Poag, C. W. *J. Foram. Res.* **9**, 85–105 (1979).
29. Wust, G. *Stratification and Circulation in the Antillean–Caribbean Basins* (Columbia University Press, New York, 1964).
30. Kowsmann, R. O. thesis, Oregon State Univ. (1973).
31. van Andel, T. H., Heath, G. R. & Moore, T. C. Jr *Geol. Soc. Am. Mem.* **143** (1975).

Evidence for the absence of biological methylation of lead in the environment

K. Reisinger, M. Stoeppler & H. W. Nürnberg

Institute of Chemistry, Institute 4 Applied Physical Chemistry,
Nuclear Research Centre (KFA), Jülich PO Box 1913,
D-5170 Jülich, FRG

Biological methylation of arsenic and mercury by pure cultures of microorganisms such as moulds[1] and bacteria[2], and by 'mixed cultures of environmental systems', such as sediments[3], sludges or canal mud[4], seem to be well known processes, although there is no unequivocal evidence of the proposed pathways[5–7]. In contrast, the biomethylation of inorganic or ionic organic lead compounds has been a topic of controversy, with no conclusive explanation of the mechanism of the reactions that might lead to the corresponding alkyl lead compounds[8–13]. We report here attempts to produce leadalkyls in anaerobic and aerobic conditions. Our study was greatly helped by our use of labelled and unlabelled substrates as methyl donors and labelled and unlabelled lead compounds as acceptors. In none of numerous experiments have we found leadalkyls to be the products of biomethylation. Experiments with organic lead salts resulted in lead tetraalkyls but only as the products of chemical reactions.

We started with the assumption that if microorganisms could really convert inorganic Pb(II) compounds and organic Pb(IV) salts, such as PbMe$_3$OAc, PbEt$_3$OAc and PbEt$_3$Cl, into the corresponding leadalkyls, a biochemical carbanion transfer reaction should occur in the bacterial cell[14–16]. We therefore used mixed cultures and an enrichment culture of methanogenic bacteria for the anaerobic experiments, and a mixed culture of aerobic bacteria and the *Escherichia coli* auxotroph 113–3 for the aerobic experiments. The enzymatically bound CH$_3$-cobalamin (Fig. 1) formed from various substrates by these organisms could act as a carbonium donor in the formation of methane and methionine, respectively, or possibly as a biological 'Grignard intermediate', transferring a carbanion (CH$_3^-$) to lead compounds, with leadalkyls as the final products. The criterion for the metabolic activity was the formation of CH$_4$/CO$_2$ and CO$_2$ by the anaerobic and aerobic bacteria, respectively. The criterion for the metabolic activity of *E. coli* 113–3 was methionine formation.

Experiments were carried out in bioreactors attached to leadalkyl collectors for trapping the alkyls blown out with the CH$_4$/CO$_2$ gas mixture produced by methanogenic bacteria, or with the air stream fed into the reactor during incubation of aerobic bacteria. The bottom of the collector was equipped with a rubber septum for sampling using syringes. Three analytical methods were applied (Fig. 2). In tracer-free experiments a gas chromatograph in combination with an atomic spectrophotometer (GC–AAS) served as the leadalkyl detector system. Samples from the collector were analysed daily by direct injection of the liquid from the collector into the gas chromatograph. Leadalkyls possibly adsorbed on the reactor wall, the bacterial mass or the sediments were also extracted after incubation and analysed by GC–AAS. Sediments and nutrient solutions were analysed for non-volatile partially biomethylated intermediates such as PbMe$_3^+$ and PbMe$_2^{2+}$ by TLC[17].

The incubation experiments with ^{14}C-labelled substrates and ^{210}Pb enabled us to use two further detection procedures. After refluxing the collector liquid to blow out ^{14}CH$_4$ and ^{14}CO$_2$ absorbed in the cooled liquid, one sample was analysed by GC–AAS and another by liquid scintillation counting. If lead had been methylated there would have been a relationship between the count rate remaining after refluxing and the peak height found by GC–AAS. In the ^{210}Pb-tracer experiments with unlabelled substrates, a multi-channel analyser was used as leadalkyl detector, and in the ^{210}Pb-tracer experiments with the labelled substrates we used the multi-channel analyser and liquid scintillation counting as detector systems. Here, too, a relationship between the ^{210}Pb and ^{14}C radioactivity would have to be found if lead compounds were methylated.

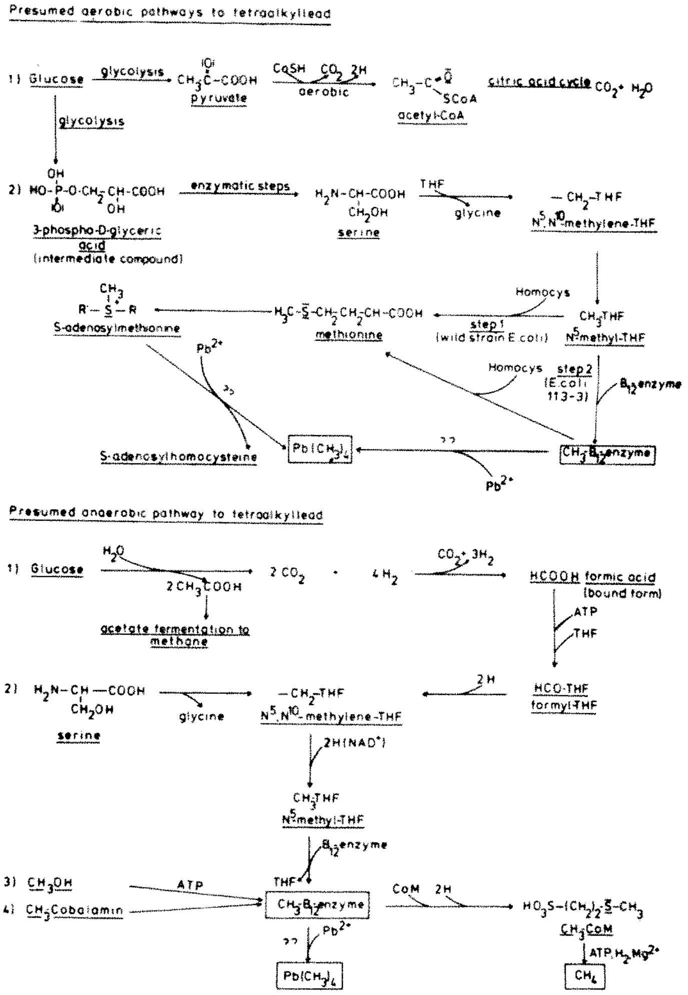


Fig. 1 The biosynthesis of the CH$_3$-B$_{12}$-enzyme from various substrates and the hypothetic methyl group transfer to lead in anaerobic and aerobic conditions.

In none of our experiments were leadalkyls or partially methylated intermediates detected as the products of biomethylation. Analysis of the collector liquid in the incubation experiments with inorganic lead and the ^{14}C-labelled substrates after refluxing revealed no Pb^{14}Me$_4$ peak, but there was a residual count rate, possibly representing a trace of Pb^{14}Me too low to be detectable by the GC–AAS. However, experiments with ^{210}Pb^{2+} as the methyl group acceptor and ^{14}C-serine and ^{14}C-methanol, as the methyl group donors, respectively, proved conclusively that this residual count rate was not due to Pb^{14}Me. As before, there was a residual ^{14}C-count rate, but no corresponding ^{210}Pb radioactivity was detectable. The incubation experiments with ^{210}Pb^{2+} and the unlabelled substrates have verified these results. At no time did the analysis of the leadalkyl trap with the multi-channel analyser exceed 135 counts per h. This count rate represented the 10^{-8} part of the count rate produced by 100 μCi ^{210}Pb radioactivity added to the culture initially.

However, incubation with PbMe$_3$OAc, PbEt$_3$OAc and the substrates summarized in Table 1 yielded PbMe$_4$ and PbEt$_4$ (Fig. 3*a*). However, as no PbMe$_3{}^{14}$Me or PbEt$_3{}^{14}$Me was found,

0028-0836/81/210228—03$01.00

Table 1 Incubation mixtures for the anaerobic and aerobic experiments at 37 °C

		Substrates								
		L-Serine		Methanol		CH_3-cobalamin		D-Glucose		Nutrient medium
Anaerobic experiments		$g\,l^{-1}$	^{14}C $\mu Ci\,l^{-1}$	$ml\,l^{-1}$	^{14}C $\mu Ci\,l^{-1}$	$g\,l^{-1}$	^{14}C $\mu Ci\,l^{-1}$	$g\,l^{-1}$	^{14}C $\mu Ci\,l^{-1}$	
River sediment +	Pb^{2+}-210	5	—	10	—	0,5	—	5	—	A
	Pb^{2+}	5	20	10	20	0,5	—	5	—	
	$PbMe_3OAc$	5	—	10	—	0,5	—	5	—	
	$PbEt_3OAc$	5	100	10	20	0,5	50	5	5	
	$PbEt_3Cl$	5	20	10	20	0,5	50	5	5	
Sewage sludge +	Pb^{2+}-210	5	20	10	—	0,5	—	5	—	A
	Pb^{2+}	5	20	10	20	0,5	—	5	—	
	$PbMe_3OAc$	5	20	10	—	0,5	—	5	—	
	$PbEt_3OAc$	5	20	10	20	0,5	50	5	5	
	$PbEt_3Cl$	5	20	10	20	0,5	50	5	5	
Methanosarcina barkeri +	Pb^{2+}-210	—	—	10	20	—	—	—	—	A
	$PbEt_3OAc$	—	—	10	20	—	—	—	—	
	$PbEt_3Cl$	—	—	10	20	—	—	—	—	
Marine sediment +	Pb^{2+}-210	—	—	10	—	—	—	—	—	
Aerobic experiments										
River sediment +	Pb^{2+}-210	—	—	—	—	—	—	5	—	B
	Pb^{2+}	5	—	—	—	—	—	5	5	
	$PbMe_3OAc$	5	—	—	—	—	—	5	5	
	$PbEt_3OAc$	5	—	—	—	—	—	5	5	
E. coli 113–3 +	Pb^{2+}	0,3	100	—	—	—	—	5	—	C
	$PbMe_3OAc$	1,0	20	—	—	—	—	5	—	
	$PbEt_3OAc$	1,0	20	—	—	—	—	5	—	

Incubation mixtures were 200g sediment or sludge, 50 ml enrichment culture or 5 ml pure *E.coli* culture, 60 mg unlabelled Pb compounds or 100 μCi $^{210}Pb(NO_3)_2$ (specific activity 1 mCi per 0.016 mg Pb), the substrates and 800 ml nutrient medium. Nutrient medium A (per 1): K_2HPO_4 (1 M), 3.7 ml; NH_4Cl (10%), 4 ml; $MgCl_2 \cdot 6\,H_2O$ (16.6%), 1 ml; $FeCl_2 \cdot 4\,H_2O$ (0.72%), 2 ml; Wolfe's minerals, 20 ml; Wolfe's vitamins, 20 ml; $NaOOCH_3 \cdot 3\,H_2O$, 1.7 g; NaOOCH, 2 g; $NaHCO_3$, 2 g; cysteine-HCl·H_2O, 0.5 g. *p*H of final mixture 7.0. Nutrient medium B (per 1): K_2HPO_4, 7 g; KH_2PO_4, 3 g; Na-citrate·(5.5 H_2O), 0.55 g; $(NH_4)_2SO_4$, 2 g; D,L-aspartic acid, 1 g; D-glucose, 5 g; $MgSO_4$, 0.05 g; NaCl, 1 g. *p*H of final mixture 7.4. Nutrient medium C (per 1): as medium B, but with vitamin B_{12}, 1 μg; D,L-homocysteine, 0.3 g.

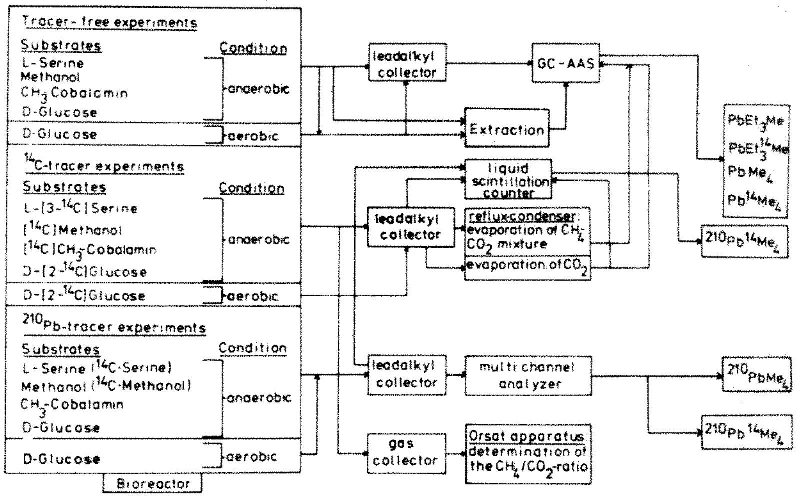


Fig. 2 Analytical methods for the determination of leadalkyls possibly formed by anaerobic or aerobic incubation in the presence of various substrates and lead compounds. During the incubation experiments leadalkyl collector used 4 ml ethanol as absorbent liquid, −30 °C; extraction of the incubation mixture was with 20 ml ligroin (b.p. 60–80 °C): GC–AAS 10 μl; detection limit 0.09 μg Pb per ml extract. Liquid scintillation counting used 0.5 ml extract and 15 ml scintillation gel; background count rate was 13 c.p.m.

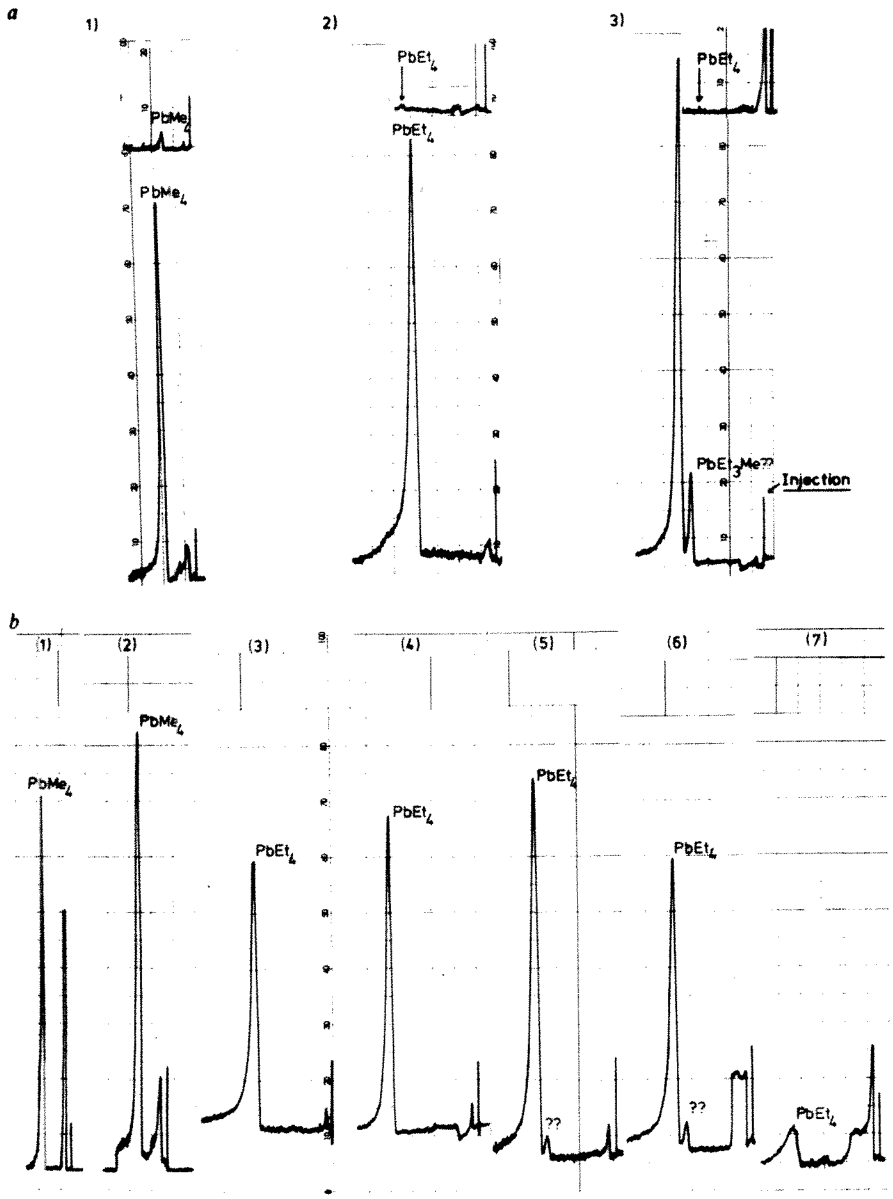


Fig. 3 The conversion of $PbMe_3OAc$, $PbEt_3OAc$ and $PbEt_3Cl$ in *in vivo* and *in vitro* conditions. *a*, Results of *in vivo* incubation of organic Pb(IV) salts with sewage sludge and L-serine substrates for 24 h. (1) Sludge + $PbMe_3OAc$; (2) sludge + $PbEt_3OAc$; (3) sludge + $PbEt_3Cl$. For comparison, blanks due to chemical side reactions in distilled water are given at the top. *b*, Results after the *in vitro* reaction (37 °C) between organic Pb(IV) salts and sulphur-containing compounds such as Na_2S, cysteine and methionine for 24 h (7: 96 h). Reaction mixture was 200 ml tridistilled water (de-aerated by boiling and gasing with N_2), 10 mg Pb compound, 10 mg sulphur compound. (1) $PbMe_3OAc$ + sulphide; (2) $PbMe_3OAc$ + cysteine; (3) $PbEt_3OAc$ + sulphide; (4) $PbEt_3OAc$ + cysteine; (5) $PbEt_3Cl$ + sulphide; (6) $PbEt_3Cl$ + cysteine; (7) $PbEt_3Cl$ + methionine.

the incubation products could not be the result of biological methyl transfer from the substrates to the corresponding ionic organic lead compounds. Moreover, tests with sulphur-containing substances have unambiguously verified the results of the important finding of Jarvie *et al.*[12] that $PbMe_4$ and naturally $PbEt_4$ are the result only of a chemical conversion induced by sulphide. Nevertheless, not only is 'free sulphide' involved in the reaction but also amino acids containing 'bound sulphide', such as cysteine and glutathione (Fig. 3*b*). On the other hand, sulphur-free amino acids, amino acids with blocked S-functions (methionine; Fig. 3*b*, chromatogram 7) or compounds with S-functions, such as thiourea, had no effect on the chemical conversion of the organic lead salts.

Incubation with $PbEt_3Cl$ (Fig. 3*a*) and $PbEt_4$ seemed to produce $PbEt_3Me$ which could be considered a product of biomethylation. However, the appearance of the presumed $PbEt_3Me$ compound even in the absence of any methylating agent, with only sulphide present (Fig. 3*b*, chromatograms 5 and 6), proved that the peak presumed $PbEt_3Me$ in Fig. 3*a* was not the result of biological methyl transfer but corresponds to an impurity. Tracer experiments have unambigiously supported these findings.

We therefore conclude that there is no basis for the assumption that biological methylation of Pb(II) occurs in the environment, whereas sulphide-induced chemical conversion of organic Pb(IV) salts into leadalkyls is possible.

We thank Dr S. Schoberth for microbiological advice and particularly K. May for support in the tracer experiments.

Received 22 December 1980; accepted 25 February 1981.

1. Challenger, F. *Chem. Rev.* **36,** 315–361 (1945).
2. Vonk, J. W. & Sijperstein, A. K. *Antonie van Leeuwenhoek* **39,** 505–513 (1973).
3. Jensen, S. & Jernelöv, A. *Nature* **223,** 753–754 (1968).
4. McBride, B. C., Merilees, H., Cullen, W. R. & Picket, W. *ACS Symp. Ser.* **82,** 94–115 (1978).
5. Wood, J. M., Kennedy, F. S. & Rosen, C. G. *Nature* **220,** 173–174 (1968).
6. Landner, L. *Nature* **230,** 452–454 (1971).
7. Bertilsson, L. & Neujahr, H. Y. *Biochemistry* **10,** 2805–2808 (1971).
8. Chau, Y. K. & Wong, P. T. S. *ACS Symp. Ser.* **82,** 39–53 (1978).
9. Huber, F., Schmidt, U. & Kirchmann, H. *ACS Symp. Ser.* **82,** 65–81 (1978).
10. Dumas, J. P., Le Roy, P., Bellonick, S., Bouchard, D. & Vaillancourt, G. *Proc. 12th Can. Symp. Water Pollution Res.* 91–100 (1977).
11. Thompson, J. A. J. & Crerar, J. A. *Mar. Pollut. Bull.* **11,** 251–253 (1980).
12. Jarvie, A. W. P., Markall, R. N. & Potter, H. R. *Nature* **255,** 217–218 (1975).
13. Craig, P. J. *Envir. Techn. Lett.* **1,** 17–20 (1980).
14. Wood, J. M. in *Lead in the Marine Environment* (eds Branica, M. & Konrad, Z.) 299–303 (Pergamon Oxford, 1980).
15. Wood, J. M. *Science* **183,** 1049–1052 (1974).
16. Wood, J. M. *et al. Fedn Proc.* **37,** 16–21 (1978).
17. Potter, H. R., Jarvie, A. W. P. & Markall, R. N. *Wat. Pollut. Control* **76,** 123–128 (1977).

Glycosaminoglycan synthesis by the apical ectodermal ridge of chick limb bud

Robert A. Kosher & Mary P. Savage

Department of Anatomy, University of Connecticut Health Center, Farmington, Connecticut 06032, USA

The apical ectodermal ridge (AER) along the distal periphery of the embryonic chick limb bud seems to have a decisive influence on the development of subjacent mesenchymal cells, preventing cartilage differentiation but ensuring the outgrowth and formation of distal limb structures[1–5]. It has been suggested that when mesenchymal cells leave the AER's influence, they undergo an intimate cell–cell interaction that triggers chondrogenic differentiation by elevating cyclic AMP levels[6–8]. Moreover, the AER may inhibit the differentiation of subjacent mesenchymal cells by causing them to secrete a considerable quantity of hyaluronate which by accumulating extracellularly may prevent the cell–cell interaction necessary to trigger differentiation[9]. Here, we have examined the synthesis of hyaluronate and other glycosaminoglycans by the AER and by limb and non-limb ectodermal tissues that do not promote limb bud outgrowth[1,5,10,11]. We find that there is a selective and substantial increase in the amount of hyaluronate produced by the AER compared with the non-inductive ectodermal tissues.

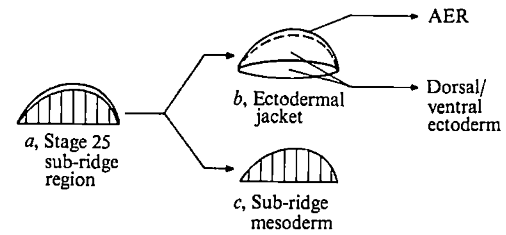


Fig. 1 Diagram of the method used to obtain AERs and dorsal/ventral limb ectoderm from the sub-ridge region of a stage 25 limb bud. Ectodermal jackets (*b*) were separated intact from the sub-ridge mesoderm (*c*) following a brief treatment with trypsin[4], and the AER was then surgically separated from the dorsal/ventral limb ectoderm by cutting with fine knives along the dotted line indicated in *b*.

Distal wing bud tips composed of the sub-ridge mesoderm capped by the AER and surrounded dorsally and ventrally by ectoderm were cut away from stage 25 embryos[12] of White Leghorn chicks (see Fig. 1*a* and ref. 4), and the ectodermal jackets (Fig. 1*b*) were separated intact from the sub-ridge mesoderm (Fig. 1*c*) following brief treatment with trypsin[4]. The AER was then surgically separated from the dorsal/ventral limb ectoderm by cutting with fine knives along the dotted line indicated in Fig. 1*b*. These procedures cleanly separate the ectodermal tissues from the sub-ridge mesoderm, and no adherence of mesodermal tissue to the isolated AERs and dorsal/ventral ectoderm is detectable with the dissecting microscope. In addition, non-limb ectodermal tissue surrounding the flank region of stage 25 embryos was removed with the aid of trypsin. The isolated AERs, the pieces of dorsal/ventral limb ectoderm and the non-limb flank ectoderm were incubated separately at 37 °C for 90 min, 17 h or 24 h in F12X medium containing 10% fetal calf serum and 100 μCi ml^{-1} of ^{3}H-glucosamine. For the 90-min incubation 150–200 AERs were used, and 30–50 AERs for the longer incubations. After incubation, glycosaminoglycans (GAGs) were extracted from the tissues and medium and the types of GAGs synthesized were determined as described in Table 1 legend.

Table 1 shows the relative amounts of the various ^{3}H-glucosamine-labelled GAGs accumulated by the AER, dorsal/ventral limb ectoderm and flank ectoderm after the three labelling periods. The only major difference in the pattern of GAG accumulation between the AER and dorsal/ventral limb ectoderm after each labelling period is in the relative amount of hyaluronate; the hyaluronate accumulated by the AER (~8%) is about twice that accumulated by the dorsal/ventral ectoderm (~4%). In contrast, there is little, if any, difference in the

Table 1. Relative amounts of the various ^{3}H-glucosamine-labelled components accumulated by the AER, dorsal/ventral limb ectoderm and flank ectoderm after various labelling periods

		Hyaluronate		Chondroitin sulphate		Other GAGs		Glycoprotein	
Labelling period	Tissue	%	Relative to AER*	%	Relative to AER*	%	Relative to AER*	%	Relative to AER*
90 min	AER	9.4±0.7	—	27.7±3.0	—	15.1±2.8	—	47.8±4.9	—
	Dorsal/ventral limb ectoderm	3.6±0.4	0.38	28.3±2.5	1.02	18.0±2.1	1.19	50.2±4.9	1.06
	Flank ectoderm	8.1±1.7	0.86	27.6±4.8	1.00	15.1±3.8	1.00	49.3±9.6	1.04
17 h	AER	7.3±1.1	—	15.0±1.8	—	14.3±3.7	—	63.6±0.6	—
	Dorsal/ventral limb ectoderm	3.8±0.7	0.52	15.2±1.3	1.01	14.4±1.3	1.01	66.7±0.8	1.05
	Flank ectoderm	4.7±2.4	0.64	10.9±2.5	0.73	17.9±2.2	1.25	66.7±2.5	1.05
24 h	AER	8.1±0.9	—	15.9±1.7	—	17.8±0.4	—	58.2±2.9	—
	Dorsal/ventral limb ectoderm	3.9±0.8	0.48	15.5±2.1	0.97	16.7±0.6	0.94	63.9±2.3	1.10
	Flank ectoderm	6.6±1.8	0.81	10.6±1.6	0.67	19.8±0.4	1.11	63.2±3.9	1.09

After the appropriate labelling period, GAGs were extracted from the tissue and medium as previously described[15]. Briefly, the tissue and medium were sonicated, and an aliquot of the sonicate removed for DNA determination[16]. The sonicates were boiled, digested with Pronase (Calbiochem), treated with ice-cold trichloroacetic acid, and the acid-soluble material was extensively dialysed to remove unincorporated isotope. The dialysates were lyophilized or evaporated to dryness, and dissolved in distilled water. The effectiveness of the isolation procedure in removing unincorporated isotope was monitored in each experiment by subjecting ^{3}H-glucosamine-containing medium which was incubated in the absence of tissue to the extraction procedure. The amount of ^{3}H-glucosamine-labelled hyaluronate was determined as previously described[17] using leech hyaluronidase (hyaluronic acid-B1, 3 hydrolase; Biometrics), an enzyme which specifically degrades hyaluronate[18,19]. Briefly, after treatment of an aliquot of the ^{3}H-glucosamine-labelled samples with the enzyme, the amount of radioactivity associated with leech hyaluronidase-sensitive and -resistant material was determined after precipitation of undegraded GAGs with cetylpyridinium chloride (CPC)[17]. Data were corrected for no enzyme control values, that is, the amount of untreated ^{3}H-glucosamine-labelled material soluble in CPC. The amount of untreated ^{3}H-glucosamine-labelled material soluble in CPC was taken to represent labelled polysaccharide chains of glycoproteins (see ref. 17). The amount of ^{3}H-glucosamine-labelled chondroitin sulphate was determined as previously described[17] using chondroitinase ABC (Miles), an enzyme which degrades the isomeric chondroitin sulphates as well as hyaluronate[20]. Aliquots of the GAG samples were treated with the enzyme, and the amount of radioactivity associated with chondroitinase-sensitive and -resistant material determined after precipitation with CPC[17]. As chondroitinase ABC degrades hyaluronate as well as the isomeric chondroitin sulphates, the amount of hyaluronate (leech hyaluronidase-sensitive GAGs) was subtracted from the amount of chrondroitinase-sensitive GAGS to determine the amount of radioactivity associated with chondroitin sulphates[17]. The other GAGs referred to represent ^{3}H-glucosamine-labelled GAGs resistant to degradation with chondroitinase ABC. About 50% of the chondroitinase ABC-resistant material produced by both the AER and dorsal/ventral limb ectoderm has been found to be sensitive to degradation by nitrous acid, and thus probably represents heparan sulphate. Values (%) represent the mean of three (90 min and 24 h) or two (17 h) determinations ± s.e.m.

* The relative amount of the appropriate ^{3}H-glucosamine-labelled component accumulated by the dorsal/ventral limb ectoderm or flank ectoderm was divided by the relative amount of that component accumulated by the AER.

0028-0836/81/210231—02$01.00

Table 2 Total amounts of the various ^{3}H-glucosamine-labelled components accumulated by the AER, dorsal/ventral limb ectoderm and flank ectoderm after various labelling periods

Labelling period	Tissue	Hyaluronate (c.p.m. per μg DNA)	Chondroitin sulphate (c.p.m. per μg DNA)	Other GAGs (c.p.m. per μg DNA)	Glycoprotein (c.p.m. per μg DNA)
90 min	AER	703	1,966	1,046	3,810
	Dorsal/ventral limb ectoderm	234	1,859	1,168	3,613
	Flank ectoderm	196	654	347	1,467
17 h	AER	8,852	18,409	18,441	79,385
	Dorsal/ventral limb ectoderm	2,241	8,411	8,332	37,616
	Flank ectoderm	1,677	3,645	5,434	20,778
24 h	AER	12,209	23,574	26,688	86,962
	Dorsal/ventral limb ectoderm	2,821	12,033	13,420	50,797
	Flank ectoderm	2,958	4,872	9,267	29,933

Values are the means of three (90 min and 24 h) or two (17 h) separate experiments.

Table 3 Amounts of the various ^{3}H-glucosamine-labelled components accumulated by the dorsal/ventral limb ectoderm and flank ectoderm relative to the amounts accumulated by the AER

Labelling period	Tissue	Hyaluronate	Chondroitin sulphate	Other GAGs	Glycoprotein
90 min	Dorsal/ventral limb ectoderm	0.37±0.08	0.98±0.06	1.17±0.17	1.01±0.11
	Flank ectoderm	0.29±0.05	0.33±0.004	0.34±0.03	0.35±0.06
17 h	Dorsal/ventral limb ectoderm	0.25±0.12	0.45±0.10	0.45± 0	0.47±0.08
	Flank ectoderm	0.18±0.17	0.19±0.13	0.30±0.01	0.25±0.08
24 h	Dorsal/ventral limb ectoderm	0.29±0.05	0.54±0.04	0.52±0.05	0.62±0.06
	Flank ectoderm	0.33±0.16	0.26±0.09	0.41±0.15	0.37±0.13

The amount of the appropriate ^{3}H-glucosamine-labelled component (c.p.m. per μg DNA) accumulated by the dorsal/ventral limb ectoderm or flank ectoderm was divided by the amount (c.p.m. per μg DNA) of that component accumulated by the AER. Values are the means of three (90 min and 24 h) or two (17 h) separate experiments ± s.e.m.

relative amounts of the other ^{3}H-glucosamine-labelled components accumulated by the two tissues. The difference in the relative amount of hyaluronate accumulated by the AER compared with the flank ectoderm is not as great as that between the AER and dorsal/ventral limb ectoderm. However, as described below, the total amount of hyaluronate accumulated by the flank ectoderm during all labelling periods is considerably less than the total amount accumulated by the AER. Also, the pattern of GAG accumulation by the isolated AERs is strikingly different from the pattern of accumulation by the mesenchymal cells directly subjacent to the AER (see ref. 9).

Table 2 shows the total amounts (c.p.m. per μg DNA) of the various ^{3}H-glucosamine-labelled GAGs accumulated by the different ectodermal tissues, and Table 3 summarizes the differences in the total amounts of the various ^{3}H-glucosamine-labelled components accumulated by the dorsal/ventral ectoderm and flank ectoderm and the AER. During a 90-min labelling period, the AER accumulates almost threefold more hyaluronate than the dorsal/ventral limb ectoderm, whereas there is little, if any, difference in the total amounts of other ^{3}H-glucosamine-labelled components accumulated by the two tissues. Thus, after a 90-min labelling period there is a selective quantitative increase in the amount of hyaluronate produced by the AER over the immediately adjacent dorsal/ventral limb ectoderm. After either a 17- or 24-h labelling period, there is an increase in the accumulation of all ^{3}H-glucosamine-labelled materials by the AER compared with the dorsal/ventral ectoderm. Note, however, that whereas the amount of chondroitin sulphate, chondroitinase-resistant GAGs and glycoprotein produced by the AER was doubled compared with the dorsal/ventral limb ectoderm, the amount of hyaluronate accumulated increased fourfold. Thus, there is a striking disproportionate increase in hyaluronate accumulation by the AER compared with the dorsal/ventral limb ectoderm. Finally, after either short (90 min) or long labelling (17 and 24 h) periods, there is a considerable reduction (three- to fivefold) in the accumulation of all ^{3}H-glucosamine-labelled components by the flank ectoderm compared with the AER (Tables 2, 3).

Obviously, our results do not prove that the substantially larger amount of hyaluronate produced by the AER than by the other ectodermal tissues is involved in the unique outgrowth-promoting effect of the AER. The possibility that the hyaluronate produced by the AER is involved in the negative differentiative effect the AER exerts on subjacent mesenchymal cells is, however, particularly intriguing. We have previously suggested that the AER may inhibit cartilage differentiation by causing subjacent mesenchymal cells to secrete large amounts of hyaluronate which prevent a cell–cell interaction necessary to trigger differentiation and/or which feed back on the cells, inhibiting their production of cartilage matrix components[9]. It is conceivable, therefore, that the hyaluronate produced by the AER may be involved in inhibiting differentiation by acting on the immediately subjacent mesenchymal cells, causing them to maintain their own high rate of hyaluronate synthesis, or perhaps by simply contributing to the hyaluronate-rich matrix surrounding the subjacent mesenchymal cells. Notably studies by Toole[13,14] indicate that exogenous hyaluronate exerts an inhibitory effect on *in vitro* chondrogenic differentiation.

This work was supported by NSF grant PCM 7925907 to R.A.K.

Received 8 October 1980; accepted 5 March 1981.

1. Saunders, J. W. in *Vertebrate Limb and Somite Morphogenesis* (eds Ede, D. A., Hinchliffe, J. R. & Balls, M.) 1–24 (Cambridge University Press, 1977).
2. Stark, R. J. & Searls, R. L. *Devl Biol.* **33,** 138–153 (1973).
3. Summerbell, D., Lewis, J. H. & Wolpert, L. *Nature new Biol.* **244,** 492–496 (1973).
4. Kosher, R. A., Savage, M. P. & Chan, S.-C. *J. Embryol. exp. Morph.* **50,** 75–97 (1979).
5. Globus, M. & Vethamany-Globus, S. *Differentiation* **6,** 91–96 (1976).
6. Kosher, R. A. & Savage, M. P. *J. Embryol. exp. Morph.* **56,** 91–105 (1980).
7. Kosher, R. A., Savage, M. P. & Chan, S.-C. *J. exp. Zool.* **209,** 221–228 (1979).
8. Ahrens, P. B., Solursh, M. & Reiter, R. S. *Devl Biol.* **60,** 69–82 (1977).
9. Kosher, R. A., Savage, M. P. & Walker, K. H. *J. Embryol. exp. Morph.* (in the press).
10. Zwilling, E. *J. exp. Zool.* **128,** 423–441 (1955).
11. Zwilling, E. *Adv. Morph.* **1,** 301–330 (1961).
12. Hamburger, F. & Hamilton, J. L. *J. Morph.* **88,** 49–92 (1951).
13. Toole, B. P. *Devl Biol.* **29,** 321–329 (1972).
14. Toole, B. P., Jackson, G. & Gross, J. *Proc. natn. Acad. Sci. U.S.A.* **69,** 1384–1386 (1972).
15. Kosher, R. A. & Lash, J. W. *Devl Biol.* **42,** 365–378 (1975).
16. Brunk, C. F., Jones, K. C. & James, T. W. *Analyt. Biochem.* **92,** 497–500 (1979).
17. Daniel, J. C., Kosher, R. A., Hamos, J. E. & Lash, J. W. *J. Cell Biol.* **63,** 843–854 (1974).
18. Meyer, K., Hoffman, P. & Linker, A. in *The Enzymes* (eds Boyer, D., Lardy, H. & Myrback, K.) 447–450 (Academic, New York, 1960).
19. Toole, B. P. in *Neuronal Recognition* (ed. Barondes, S. H.) 275–329 (Plenum, New York, 1976).
20. Saito, H., Yamagata, T. & Suzuki, S. *J. biol. Chem.* **243,** 1536–1542 (1968).

Induction of anti-H-2 antibodies without alloantigen exposure by *in vivo* administration of anti-idiotype

Jeffrey A. Bluestone, Susan O. Sharrow, Suzanne L. Epstein, Keiko Ozato & David H. Sachs

Transplantation Biology Section, Immunology Branch, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20205, USA

Anti-idiotypic reagents have been used recently to manipulate both idiotype (ID) expression and antigen-binding antibody activity of immunized animals[1–11]. The *in vivo* administration of anti-idiotypic antibodies followed by antigen exposure has usually resulted in suppression of ID-positive antigen binding molecules[6–8], although, in other circumstances, increased idiotype expression and antigen-binding activity has also been observed[1–5]. Extension of these findings to the immune response to transplantation antigens is of obvious practical and theoretical importance. There have been several reports describing the production of anti-idiotypic reagents against alloantisera[9–11], although the generation of anti-ID antibodies in this way has generally been difficult to reproduce[11]. Therefore, in the study reported here we used xenogeneic anti-idiotypic antisera raised against monoclonal anti-H-2K^k antibodies which offer the advantage of reproducibility of anti-idiotype production[12]. These anti-idiotypic reagents were produced by immunizing swine and rabbits with either purified 3-83P[13] or 11-4.1 (ref. 14), both of which are IgG2a, k anti-H-2K^k antibodies. We found that none of these reagents detected significant levels of idiotype in anti-H-2 antisera produced against H-2K^k antigens by conventional means. However, BALB/c mice treated with these anti-idiotypes consistently produced large amounts of idiotype, and in some animals this induced response included idiotype-positive antibodies which bound to H-2K^k antigens. Thus, treatment with anti-idiotype has been found to induce anti-H-2K^k antibodies in the absence of exposure to the H-2K^k antigen.

Immune antisera were absorbed with LPC-1 myeloma (IgG2a, k) coupled to Sepharose to remove all anti-normal mouse immunoglobulin. Purified anti-3-83P or anti-11-4.1 antibodies were obtained by specific adsorption to and elution from Sepharose columns bearing the appropriate idiotype. We have previously demonstrated that these purified anti-idiotypic reagents detect unique idiotypic determinants on the monoclonal antibodies but do not cross-react with each other or with a series of other monoclonal anti-H-2K^k antibodies, some of which share identical specificities[12]. The effects of *in vivo* administration of these anti-ID reagents on both the expression of ID and the generation of H-2K^k antigen binding molecules were examined.

Table 1 Idiotype expression on serum antibodies from BALB/c mice immunized with anti-idiotypic antibodies or alloantigen

		HAI titre ($\log_2$) of sheep red blood cells coated with:			
		3-83P ID		11-4.1 ID	
		Pig α	Rabbit α	Pig α	Rabbit α
Group	Treatment	3-83P	3-83P	11-4.1	11-4.1
(1)	BALB.K skin graft and spleen cells	≤1	≤1	≤1	≤1
(2)	Rabbit anti-3-83P antibodies	8–9	6–8	1–2	≤1
(3)	Pig anti-11-4.1 antibodies	1–2	≤1	5–6	5–6
(4)	Rabbit anti-normal mouse immunoglobulin	≤1	≤1	≤1	≤1
(5)	Pig anti-normal mouse immunoglobulin	1–2	<1	1–2	<1

In group 1, BALB/c mice were immunized against H-2^k antigens by a BALB.K tail skin graft and subsequent boosting i.p. with 2×10^7 BALB.K spleen cells per mouse. Other groups of BALB/c mice were treated on day 0 and 3 with 20 μg of the antibodies injected i.p. in saline. On day 42, mice were bled and separate pools of sera from each group of mice were assayed for both idiotype expression and for anti-H-2K^k activity (Fig. 1). Idiotype levels were determined using a haemagglutination inhibition assay (HAI) as described elsewhere[4].

In the first experiment (Table 1) groups of BALB/c mice were treated with alloantigen (C3H skin graft and lymphoid cells), anti-idiotypic antibodies, or anti-normal mouse immunoglobulin (eluted from LPC-1 coupled to Sepharose). The animals were bled on day 42 and a serum pool from each group tested for idiotype expression in a haemagglutination inhibition assay (HAI; ref. 15). The pool of anti-H-2K^k alloantiserum (cytotoxic titre 1:400) did not express detectable levels of ID, which confirmed previous results on pools of conventional anti-H-2^k antisera made in BALB/c and other strains of mice[12]. In contrast, treatment with purified anti-idiotypic antibodies (20 μg intraperitoneally (i.p.), days 0 and 3) resulted in the

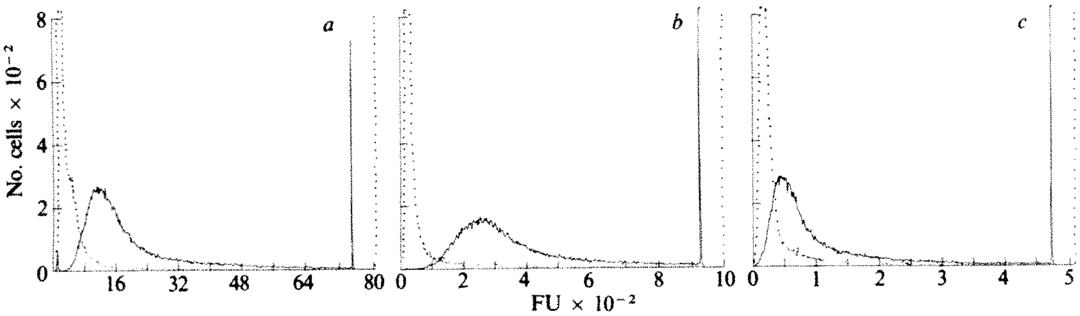


Fig. 1 Antiserum from alloimmunized and anti-idiotype treated mice were examined for anti-H-2K^k activity by flow cytofluorometry. Lymph node cells from C3H (H-2K^k) or C3H.OH (H-2K^d) mice were stained with test serum from individual mice of groups 1–4 (Table 1), counterstained with a mixture of fluoresceinated goat $F(ab')_2$ anti-mouse IgG1 (1:32) and fluoresceinated goat $F(ab')_2$ anti-mouse IgG2 (1:32) and analysed in a FACS II (Becton Dickinson). Fluorescence data were collected on 5×10^4 viable cells and displayed as a 1,000-channel cell frequency histogram in which fluorescence intensity was plotted (abscissa) against cell number (ordinate). Fluorescence intensity was converted to fluorescence units (FU; ref. 17) for comparison. In *a*, sera from a pool of mice immunized with a C3H tail skin graft and lymphoid cells were examined for anti-H-2^k activity on lymph node cells from C3H (solid line) and C3H.OH (dotted line) mice. *b*, Fluorescence profile of the binding of anti-H-2K^k antibodies in an Erhlich's induced ascites from mouse 1529, immunized on day 0 and 3 with 20 μg of purified rabbit anti-3-83P antibodies and boosted with 20 μg of anti-idiotype 1 week before being given the ascites tumours (group 2, Table 1). *c*, Profile of anti-H-2K^k binding of an ascites from mouse 1505 similarly treated with pig anti-11-4.1 anti-idiotypic antibodies (Group 3, Table 1).

induction of idiotype-positive (ID$^+$) molecules in the serum. The observed inhibition of haemagglutination by immune sera from anti-idiotype treated mice was not due to an anti-pig or anti-rabbit immunoglobulin response because: (1) idiotype-positive molecules induced in mice treated with pig anti-11-4.1 antibodies were detected by both rabbit and pig anti-idiotype agglutinators (Table 1), and (2) BALB/c mice treated with either pig or rabbit anti-normal mouse immunoglobulin produced no significant levels of idiotype-positive molecules. ID$^+$ molecules were detected as early as 1 week after anti-idiotype administration, reached a plateau at ~3 weeks, and remained detectable for several months after immunization. Only the homologous ID was expressed after *in vivo* treatment with anti-idiotype. For example, in group 2, the injection of anti-3-83P anti-idiotypic antibodies resulted in the production of 3-83P ID$^+$ but not 11-4.1 ID$^+$ molecules.

The ascites from anti-idiotype treated mice were next examined for anti-H-2K^k binding activity by flow cytofluorometry (Fig. 1). Ascites from individual immunized animals were incubated with C3H (H-2K^k) or C3H.OH (H-2K^d) lymph node cells. Fluoresceinated F(ab')$_2$ anti-IgG1 and F(ab')$_2$ anti-IgG2 antibodies were then added to the antibody-coated cells and the stained cells analysed using the fluorescence activated cell sorter (FACS II). Figure 1 illustrates the anti-H-2K^k activity of two mice, one from group 2 (1529) and one from group 3 (1505). Anti-H-2K^k binding activity was detected in three of six mice treated with rabbit anti-3-83P anti-idiotype and in two of five mice treated with pig anti-11-4.1. None of 10 control mice treated with either rabbit or pig anti-normal mouse immunoglobulin produced detectable levels of anti-H-2K^k antibodies (data not shown). The specificity of the antigen-binding molecules produced by the anti-idiotype treated mice was confirmed using H-2 congeneic recombinant strains of mice differing only at the H-2K^k region[16]. The anti-H-2 antibodies induced in these animals bound to B10.AKM (H-2K^k) but not B10.MBR (H-2K^b) lymph node cells.

To assess idiotype expression on these anti-idiotype induced anti-H-2K^k antibodies, we examined the ability of purified anti-idiotypic antibodies or anti-normal mouse immunoglobulin to block antibody binding to C3H lymph node cells. As seen in Fig. 2, anti-normal mouse immunoglobulin had no effect on the binding to the appropriate cells of either conventional allo-

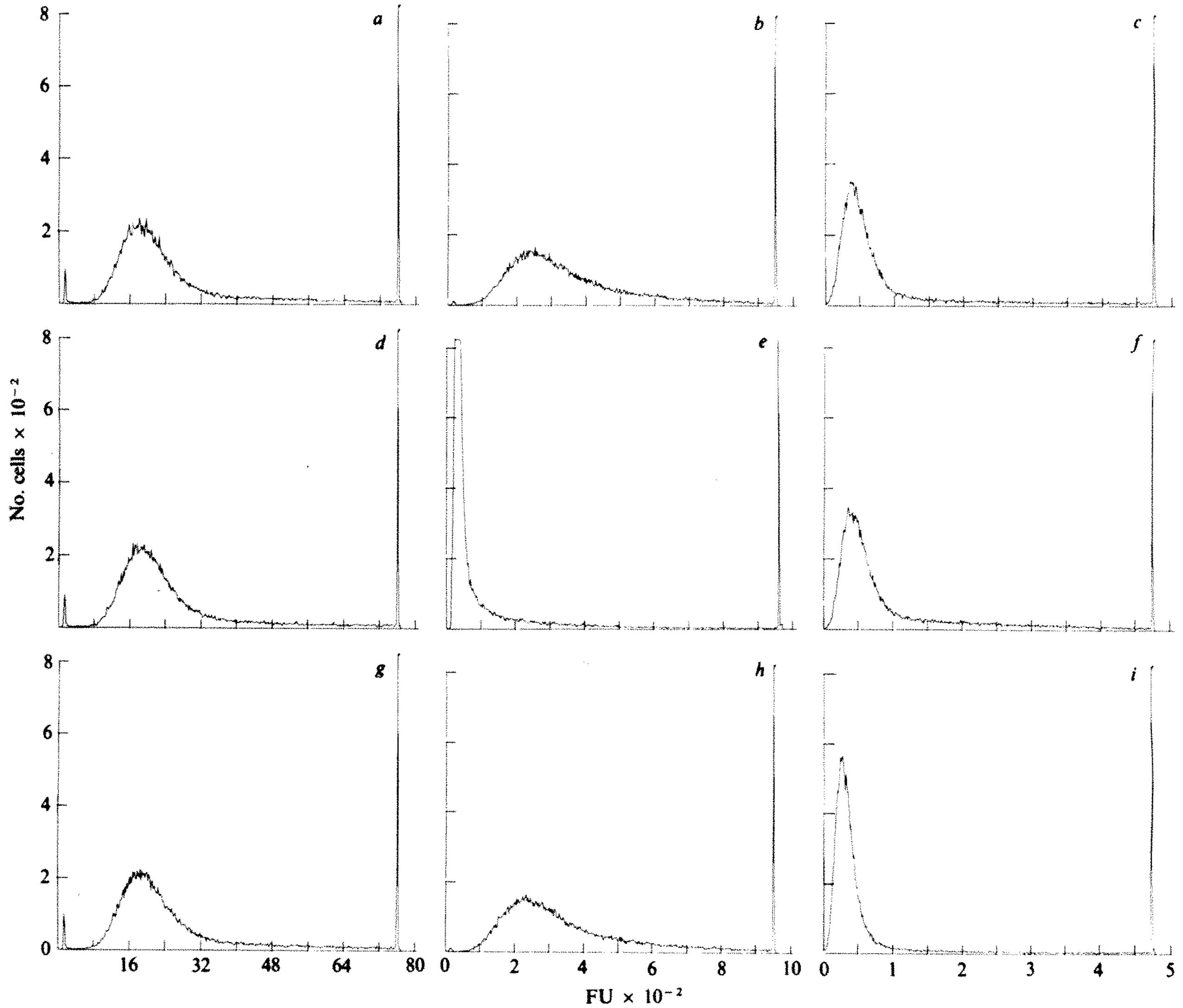


Fig. 2 Anti-H-2K^k antibodies from alloimmunized and anti-idiotype treated mice were examined for idiotype expression by a binding inhibition assay. Serum (25 μl) from treated mice was mixed with either 25 μl of anti-idiotype (100 μg ml^{-1}) or 25 μl of anti-normal pig immunoglobulin (100 μg ml^{-1}). After incubation for 1 h, the test sera plus inhibitor were incubated with lymph node C3H cells and binding assessed as described in Fig. 1 legend. Binding of conventional pooled alloantisera to C3H cells was not appreciably inhibited by pig anti-normal mouse IgG (*a*), pig anti-3-83P (*d*), or pig anti-11-4.1 (*g*). Columns 2 (profiles *b*, *e*, and *h*) and 3 (profiles *c*, *f* and *i*) illustrate the inhibitory ability of the same pig antibodies on the binding of anti-H-2K^k antibodies induced in mice treated with rabbit anti-3-83P (1529) or pig anti-11-4.1 (1505), respectively. As the fluorescence units (FU) in each set of profiles differ, the level of inhibition must be compared with the reactivity of the serum with the congeneic H-2K^d cells (Fig. 1). Such comparisons indicate that <10% of the anti-H-2K^k binding activity in mice treated with anti-idiotype remained after inhibition with the appropriate anti-idiotype (profiles *e* and *i*).

antibodies or anti-idiotype induced anti-H–2K^k antibodies. In addition, antigen binding of pooled, high-titre, conventionally raised anti-H–2K^k antibodies was relatively unaffected by either anti-3-83P or anti-11-4.1 anti-idiotypic antibodies.

However, anti-idiotypic antibodies completely inhibited the binding of anti-H–2K^k antibodies induced by homologous anti-ID, as shown by a shift in intensity of the labelled cells to a level equivalent to that of controls not preincubated with alloantibodies. Specifically, purified pig anti-3-83P antibodies completely inhibited the H–2K^k antigen binding activity of antibodies induced in mouse 1529 by rabbit anti-3-83P. In contrast, anti-11-4.1 antibodies had no effect on the binding of serum antibodies from mouse 1529 but significantly inhibited the binding of antibodies induced in mouse 1505 by pig anti-11-4.1 anti-idiotype. These results suggest that the anti-H–2K^k antibodies induced in anti-idiotype treated mice expressed idiotopes closely related to those of the original hybridoma anti-H–2K^k reagents. However, although both the original monoclonal anti-H–2K^k antibodies were of the IgG2a subclass, the idiotypic determinants were found to be expressed on immunoglobulins of both the IgG1 and IgG2 subclasses by successive adsorption to and elution from specific goat anti-mouse IgG1 and anti-mouse IgG2 coupled to Sepharose (data not shown).

The finding that anti-idiotypic reagents produced against each monoclonal anti-H–2K^k antibody recognized determinants detected on few alloantibodies or other monoclonal anti-H–2K^k antibodies of similar specificity is most consistent with a large repertoire of anti-H–2 antibodies binding a given allodeterminant. Thus, the normal immune response to H–2 antigens such as H–2K^k may include a large number of individual clones each of which expresses private, or non-cross-reactive idiotypic determinants. Such polyclonality may be the reason that production of anti-idiotypic sera against pooled alloantisera has been so difficult[11]. Despite such polyclonality, these studies have demonstrated that among the idiotype-positive molecules induced in anti-ID treated mice are molecules which bind specifically to H–2K^k antigens. The anti-H–2K^k activity induced represents expression of V region genes which are common to BALB/c mice but which are only occasionally used in the normal response to H–2K^k. Therefore, these findings are consistent with a possible sharing of variable region structures in the anti-H–2K^k monoclonal antibodies with those of alloantibodies induced by anti-idiotype.

The induction of antigen-binding ID$^+$ molecules in the absence of exposure to antigen does not usually occur in systems involving soluble antigens after treatment with heterologous anti-ID[1,2,4]. This difference may depend in part on the increased sensitivity of detection of antigen binding activity allowed by FACS II. However, these findings may also reflect an intrinsic difference in the response to histocompatability antigens compared with most other antigens[9,10]. The ability of lymphoid cells to recognize H–2 antigens is of central importance in the immune response. Anti-idiotypic antibodies to monoclonal anti-H–2 antibodies may be useful as a probe for analysing the cellular interactions involved in immune responses and as a new approach for the specific manipulation of these responses.

We thank Marion Knode and Nanci Mayer for technical assistance in the production and purification of the monoclonal antibodies and anti-idiotype.

Received 2 December 1980; accepted 4 March 1981.

1. Cazenave, P. A. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5122 (1977).
2. Urbain, J., Winkler, M., Franssen, J. P. & Colligan, C. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5126 (1977).
3. Trenkner, E. & Riblet, R. *J. exp. Med.* **142,** 1121 (1975).
4. Sachs, D. H., El-Gamil, M. & Miller, G. *Eur. J. Immun.* (in the press).
5. Kelsoe, G., Reth, M. & Rajewsky, K. *Immun. Rev.* **52,** 75 (1980).
6. Hart, D. A., Wang, A. C., Pawlak, L. & Nisonoff, A. *J. exp. Med.* **135,** 1293 (1972).
7. Eichmann, K. *Eur. J. Immun.* **4,** 296 (1974).
8. Cozenza, H. & Kohler, H. *Proc. natn. Acad. Sci. U.S.A.* **69,** 2701 (1972).
9. Ramseier, H. *Expl Cell Biol.* **47,** 107 (1979).
10. Frischnecht, H., Binz, H. & Wigzell, H. *J. exp. Med.* **147,** 500 (1978).
11. Krammer, P. H. *Curr. Topics Microbiol. Immun.* (in the press).
12. Sachs, D. H., Bluestone, J. A., Epstein, S. L. & Ozato, K. *Trans. Proc.* **13,** 953 (1981).
13. Ozato, K., Mayer, N. & Sachs, D. H. *J. Immun.* **124,** 533 (1980).
14. Oi, V. T., Jones, P. P., Goding, L. A., Herzenberg, L. A. & Herzenberg, L. A. *Curr. Topics Microbiol. Immun.* **81,** 115 (1978).
15. Fathman, C. G. & Sachs, D. H. *J. Immun.* **116,** 959 (1976).
16. Sachs, D. H., Arn, J. S. & Hansen, T. H. *J. Immun.* **123,** 1965 (1979).
17. Sharrow, S. O., Ozato, K. & Sachs, D. H. *J. Immun.* (in the press).

Cell interactions modulate embryonal carcinoma cell differentiation into parietal or visceral endoderm

Brigid L. M. Hogan & Amanda Taylor

Imperial Cancer Research Fund, Mill Hill Laboratories,
Burtonhole Lane, London NW7 1AD, UK

Eileen Adamson

Department of Zoology, University of Oxford, South Parks Road,
Oxford OX1 2JD, UK

F9 is a clonal line of mouse teratocarcinoma-derived embryonal carcinoma (EC) cells which shows very little spontaneous differentiation *in vivo* or *in vitro*[1]. Recently, however, it was reported[2–6] that F9 monolayers treated with retinoic acid and dibutyryl cyclic AMP differentiate into an early embryonic cell type known as parietal endoderm, which is one of two distinct populations of extra-embryonic endoderm that differentiate in the normal mouse embryo shortly after implantation[7,8]. The other population is the visceral endoderm, and the two differ not only in their morphology and location within the embryo, but also in their biochemical properties[9–12]. Parietal endoderm cells, for example, do not synthesize α-fetoprotein[13] (AFP), whereas visceral endoderm cells do[13]. There is evidence[8] that during embryogenesis parietal and visceral endoderm are derived from a common precursor population known as the primary endoderm, and recent experiments with cultured mouse embryos[14] suggest that the phenotype of these cells can be modulated by contact with different embryonic tissues. We now show that if F9 EC cultures are treated with retinoic acid when they are in the form of small aggregates, they differentiate on the outer surface cells which morphologically resemble visceral rather than parietal endoderm. In addition, the cells synthesize and secrete AFP, identified by immunoprecipitation and immunoperoxidase reactions using specific anti-AFP immunoglobulin. One interpretation of this result is that F9 cells treated with retinoic acid differentiate first into multipotent cells analogous to the primary endoderm of the normal embryo which then express either the mature parietal or visceral phenotype depending on the nature of the intercellular contact signals they receive. We therefore believe that F9 EC cells may be even more useful than previously supposed for biochemical studies on factors controlling gene expression during mammalian embryogenesis.

Small clumps of 30–50 EC cells are grown in suspension in bacteriological Petri dishes in medium containing 5×10^{-8}M retinoic acid. After 2–4 days most aggregates resemble simple embryoid bodies[15,16] with an outer layer of endoderm surrounding a core of undifferentiated cells. These aggregates often become hollow, and for a while resemble cystic embryoid bodies[15] (Fig. 1*a*). Later, however, this organization breaks down as an increasing proportion of the inner cells develop an endoderm-like morphology, and internal tissues analogous to embryonic ectoderm and mesoderm have never been observed, even after prolonged incubation. When aggregates cultured for 7 days are examined in the electron microscope, several kinds of endoderm cell are found. As shown in Fig. 1*c* and *d*, most

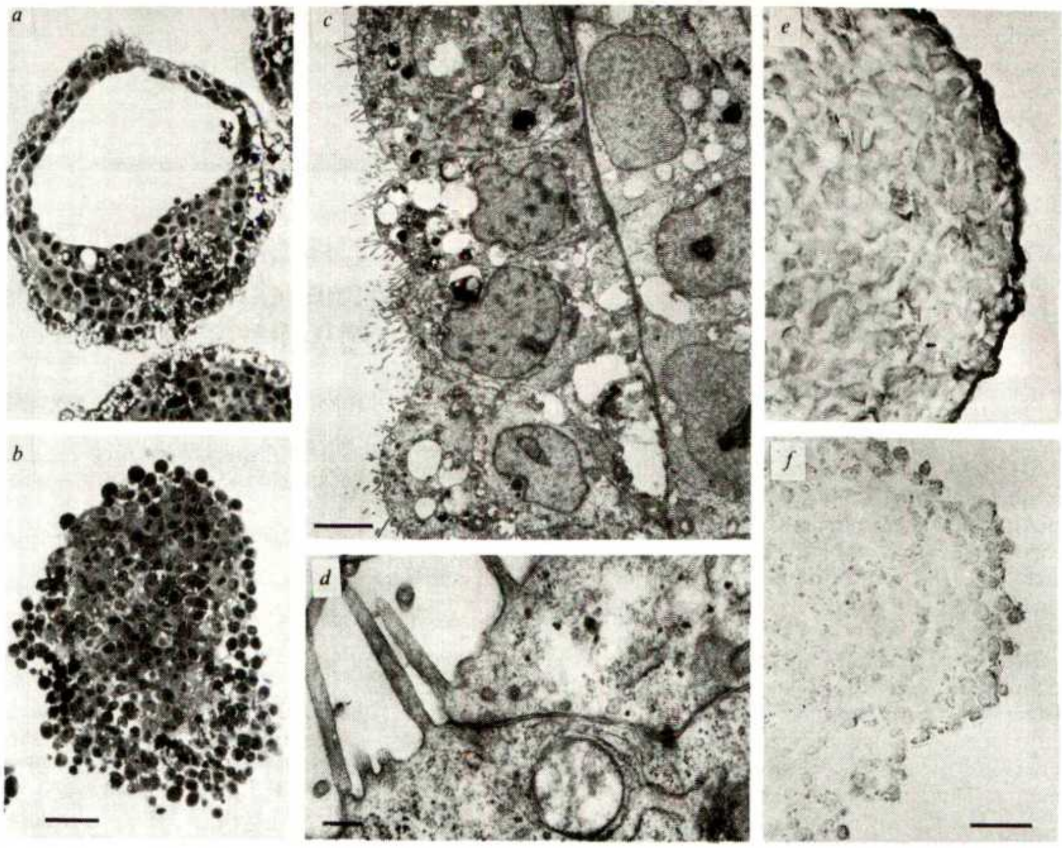


Fig. 1 Morphology of F9 cell aggregates treated with retinoic acid. *a*, Section through an aggregate of F9 cells cultured in the presence of 5×10^{-8} M retinoic acid for 7 days. Endoderm cells form an epithelial layer around the hollow core of still undifferentiated embryonal carcinoma cells. *b*, Section through an aggregate of F9 cells cultured in the presence of 1×10^{-7} retinoic acid and 10^{-3} M dibutyryl cyclic AMP for 5 days. In these conditions the endoderm cells do not form an epithelial layer but are rounded and loosely packed, with few intercellular junctional complexes. In the electron microscope all the differentiated cells have a typical parietal endoderm morphology. Scale bar, 50 μm for both *a* and *b*. *c*, Electron micrograph of visceral endoderm-like cells present in an aggregate of F9 cells exposed to 5×10^{-8} M retinoic acid for 7 days. The cells are joined apically by junctional complexes and have numerous microvilli, endocytotic vacuoles, lysosomes and lipid droplets. These features are all characteristic of visceral endoderm of the normal mouse embryo[7,14,17]. Scale bar, 5 μm; *d*, Detail of a desmosome and an apical junctional complex between two adjacent visceral endoderm-like cells in an aggregate exposed to 5×10^{-8} M retinoic acid for 7 days. Such intercellular connections are not found between parietal endoderm cells in the normal embryo[7,17]. Scale bar, 0.2 μm. *e*, *f*, Immunoperoxidase staining for AFP on sections of F9 aggregates. *e*, Cells were cultured for 5 days in the presence of 5×10^{-8} M retinoic acid in the standard culture medium and then for 2 days in serum-free medium[19]. This was to avoid retention of calf serum proteins which might cross-react with rabbit anti-mouse AFP. The antibody was also pre-absorbed with Sepharose-linked calf serum proteins to reduce further the possibility of staining calf AFP which could have been endocytosed by the F9 cells. The aggregates were fixed in acidified ethanol and embedded, sectioned and reacted as described elsewhere[13]. *f*, Aggregates cultured in standard medium containing 1×10^{-7} M retinoic acid and 10^{-3} dibutyryl cyclic AMP for 5 days were fixed, embedded, sectioned and reacted with rabbit anti-mouse AFP as described above, except that the antiserum was not pre-absorbed with calf serum proteins. No intracellular peroxidase staining is present. Scale bars, 20 μm *e* and *f*.

resemble visceral endoderm of the post-implantation mouse embryo[7,12,14,17] and have numerous microvilli and pinocytotic and lysosomal vacuoles, a thin endoplasmic reticulum and prominent intercellular apical junctional complexes. Others resemble parietal endoderm[7,12,17], with an extensive endoplasmic reticulum engorged with secretory material and few microvilli or intercellular junctional complexes. These cells are usually associated with pools of basement membrane material. The remainder have an intermediate morphology. Control aggregates incubated without retinoic acid never develop a continuous outer layer of endoderm, although some produce small 'blebs' of visceral endoderm-like cells.

As cells with a visceral endoderm morphology are present in aggregates incubated with retinoic acid we looked for the synthesis and secretion of the visceral endoderm marker protein, AFP. Aggregates were incubated with ^{35}S-methionine, and AFP immunoprecipitated from the culture medium with specific antiserum[13]. Synthesis of AFP could be detected after only 3 days' exposure to retinoic acid, and increased up to 7 days (Fig. 2*a*). At this time AFP represented about 0.05% of the total newly synthesized protein secreted by F9 aggregates, compared with a value of 4% for visceral yolk sacs isolated from 11.5-day embryos, although the comparison is only approximate because of the very different cellular composition and pool sizes of yolk sacs and aggregates. Figure 2*a* shows that the radioactive polypeptide immunoprecipitated by anti-AFP antiserum from the medium of F9 aggregates treated with retinoic acid co-migrates on SDS-polyacrylamide gels with AFP synthesized by visceral endoderm isolated from 13.5-day mouse embryos. Control F9 cell aggregates not exposed to retinoic acid synthesized very low and variable amounts of AFP (Fig. 2*a*), probably as a result of some spontaneous differentiation into endoderm. Experiments in which the concentration of retinoic acid was varied over 10^{-8}–10^{-6}M showed that the optimal concentration for inducing AFP production was 5×10^{-8}M. When aggregates exposed to retinoic acid for 7 days were sectioned and stained for AFP by the indirect immunoperoxidase technique using specific antiserum[13], a positive reaction was obtained in most endoderm cells (Fig. 1*e*). Control aggregates not exposed to retinoic acid had no intracellular staining, except where a bleb of endoderm cells was present (data not shown). Because the rabbit anti-mouse AFP antiserum was pre-absorbed with calf serum proteins and did not react with calf AFP in immunodiffusion assays, we are confident that the stain is identifying endogenously synthesized AFP. However, we cannot exclude the possibility that the AFP is synthesized and secreted by only a minority of the visceral endoderm cells and is then endocytosed from the medium by the remainder.

Earlier, Strickland and Sawey[4] failed to find evidence for AFP synthesis by F9 cells in monolayer cultures. In their experiments the cells were exposed to 1×10^{-7} M retinoic acid for 4 days. We have treated monolayers with 5×10^{-8} M retinoic acid, and observe significant AFP synthesis starting at 7 days, by which time the cultures are very dense and localized clumping between cells with endoderm and EC-like morphology has developed (Fig. 2*c*). In agreement with the observations of Strickland *et al.*[6], we find that F9 monolayers exposed to both retinoic acid and dibutyryl cyclic AMP differentiate into endoderm cells

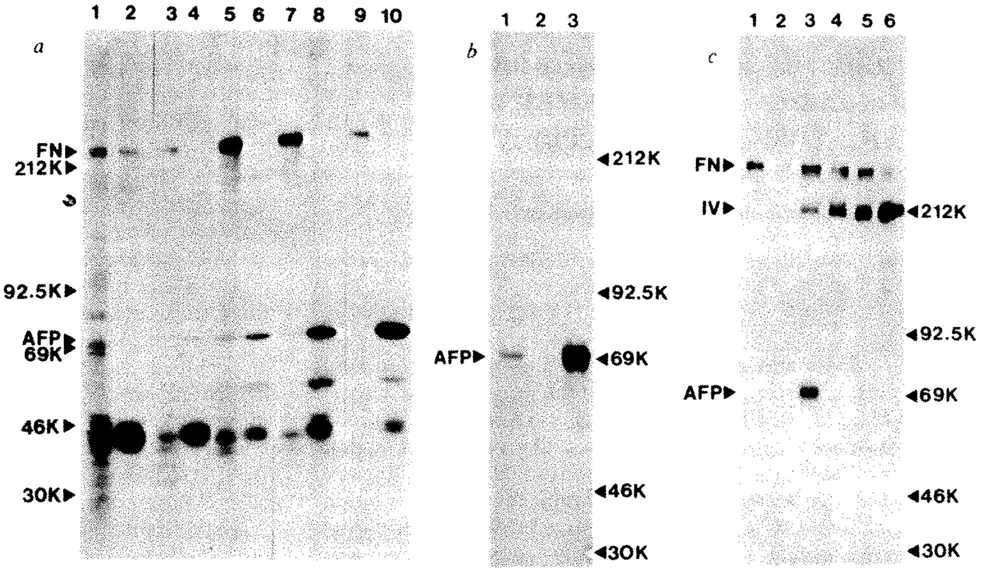


Fig. 2 Synthesis and secretion of AFP by F9 cells treated with retinoic acid. *a*, Time course of AFP synthesis by aggregates cultured in retinoic acid. Exponential cultures of F9 EC cells (from a line supplied by Colin Stewart, which had been recloned after *in vivo* passage as a tumour) were briefly trypsinized, and small aggregates resuspended in bacteriological Petri dishes in Dulbecco's modified Eagle's medium (DMEM) containing 10% fetal bovine serum (FBS) and 5×10^{-8} M retinoic acid (Sigma, added from a 10^{-4} M stock solution in alcohol and protected from light). Control aggregates were cultured without retinoic acid. The medium was changed every 48 h. After 3 days ~100 aggregates were removed every 24 h, washed twice in DMEM containing 1 μg ml^{-1} methionine and 10% dialysed FBS (labelling medium) and incubated for ~11 h in 1 ml labelling medium with 100μCi ^{35}S-methionine (specific activity 1,000 Ci mmol^{-1}; Radiochemical Centre), with and without retinoic acid. After incubation the aggregates were removed by sedimentation and 400-μl aliquots of the medium used for immunoprecipitation. Samples were diluted into 1 ml of 0.4 M NaCl, 0.005 M EDTA, 0.05 M Tris (*p*H 8.0) and 1% v/v NP40 (NET buffer) and 10 μl of rabbit anti-mouse AFP antiserum[13] added. After 1.5 h at room temperature 50 μl of protein A–Sepharose (Pharmacia, resuspended in NET buffer) were added and the incubation continued with continuous mixing at 4 °C for 1.5 h. The beads were then washed by centrifugation twice in NET buffer, once in 0.01 M Tris, *p*H 6.8, and finally boiled for 3 min in 30 μl gel sample buffer[18] (0.0625 M Tris, *p*H 6.8, 2% SDS, 10% glycerol, 5% β- mercaptoethanol). SDS-polyacrylamide gel electrophoresis was carried out according to the method of Laemmli[18], using 5–10% gradient slab gels. ^{14}C-labelled marker proteins (Radiochemical Centre) were included in the gel and their position is shown on the left-hand side of the figure. After electrophoresis the gels were fixed, impregnated with EN^{3}HANCE (NEN), dried and exposed to X-ray film (Kodak SB-5) at −70 °C. Tracks 1, 3, 5, 7 and 9, proteins immunoprecipitated with anti AFP antiserum from medium of control aggregates cultured for 3, 4, 5, 6 and 7 days in the absence of retinoic acid. The medium samples contained 1.59×10^6, 0.37×10^6, 0.41×10^6, 0.46×10^6 and 0.54×10^6 c.p.m. in total, acid-precipitable protein, respectively. Tracks 2, 4, 6, 8 and 10, proteins immunoprecipitated from medium of aggregates cultured for 3–7 days in the presence of 5×10^{-8} M retinoic acid. The medium contained 1.15×10^6, 0.42×10^6, 0.30×10^6, 0.30×10^6 and 0.53×10^6 c.p.m. in total acid-precipitable protein, respectively. Newly synthesized AFP migrates as a radioactive polypeptide of molecular weight (MW) ~70,000. The polypeptide of MW ~57,000 probably represents an AFP degradation product (E.A., unpublished observation), while the identity of the diffuse band at MW 46,000 is unknown. This band was not observed when the antiserum was pre-absorbed with adult mouse serum (see *b* and *c*). At all times, some radioactive fibronectin MW ~265,000 is nonspecifically bound to the protein A–Sepharose. From this, it can be seen indirectly that the amount of fibronectin synthesized and/or released into the culture medium by F9 cells is reduced after exposure to retinoic acid. Similar results to those shown were obtained with a line of F9 cells obtained from D. Solter. *b*, Co-migration of AFP synthesized by F9 aggregates and by normal visceral endoderm. After 6 days in culture in 5×10^{-8} M retinoic acid, aggregates were incubated for 12 h in labelling medium containing 200 μCi ml^{-1} ^{35}S-methionine and 5×10^{-8} M retinoic acid. Four visceral yolk sacs dissected from 13.5-day C3H/He mouse embryos were incubated for 18 h in labelling medium containing 100 μCi ml^{-1} ^{35}S-methionine. Aliquots of the culture medium were immunoprecipitated with 4 μg of IgG prepared from rabbit anti-mouse AFP antiserum pre-absorbed with adult mouse serum proteins. Track 1, immunoprecipitate of medium from F9 aggregates; track 2, immunoprecipitate of medium from visceral yolk sacs using 10 μl of pre-immune rabbit serum; track 3, immunoprecipitate of medium from visceral yolk sacs. *c*, Synthesis of AFP by F9 monolayers incubated with and without retinoic acid and dibutyryl cyclic AMP for 7 days. 1×10^4 F9 cells were seeded into 35-mm tissue culture dishes in culture medium with and without 5×10^{-8} M retinoic acid and 10^{-3} M dibutyryl cyclic AMP. The medium was changed every 48 h and after 7 days replaced with 1 ml of labelling medium containing 100 μCi ml^{-1} ^{35}S-methionine. After 13.5 h incubation the medium was removed and the cells trypsinized and counted. Cell numbers per dish were as follows: control, 5.63×10^6; with retinoic acid, 1.08×10^6; with retinoic acid and dibutyryl cyclic, 1.71×10^6. The volume of culture medium used for the immunoprecipitation was adjusted to take into account the difference in cell number. Total c.p.m. in acid-precipitable protein in the medium samples were as follows: control, 6.87×10^4 (in 58 μl); with retinoic acid, 6.56×10^5 (in 300 μl); with retinoic acid and dibutyryl cyclic AMP, 1.47×10^6 (in 189 μl). Medium was immunoprecipitated with either 4 μg rabbit anti-mouse AFP IgG (see *b*) or 10 μl of non-immune rabbit serum (NRS). Track 1, control medium immunoprecipitated with anti-AFP IgG; track 2, control medium immunoprecipitated with NRS; track 3, medium from cells treated with retinoic acid, immunoprecipitated with anti-AFP IgG; track 4, same medium immunoprecipitated with NRS; track 5, medium from cells treated with retinoic acid and dibutyryl cyclic AMP, immunoprecipitated with anti-AFP IgG; track 6, same medium immunoprecipitated with NRS. In addition to the nonspecific precipitation of fibronectin, tracks 3–6 show the nonspecific precipitation of type IV procollagen (MW ~190,000) synthesized by F9 cells in response to retinoic acid.

which synthesized large amounts of laminin and type IV procollagen and no fibronectin (unpublished data). As shown in Fig. 2*c*, these cells do not synthesize AFP, a result which supports their classification as exclusively parietal endoderm. F9 cells exposed to retinoic acid and dibutyryl cyclic AMP in aggregates (Fig. 1*b*) also develop a parietal endoderm morphology as judged by light and electron microscopy, and do not synthesize AFP or stain positively with anti-AFP antiserum using the indirect immunoperoxidase technique (Fig. 1*f*).

One interpretation of our results, and those of Strickland *et al.*[4,6], is that F9 cells exposed to retinoic acid first differentiate into endoderm cells which have the potential to express one of two alternative phenotypes. In the absence of intercelluar contacts with undifferentiated cells (for example, in low density monolayer cultures) or in the presence of high levels of cyclic AMP, these endoderm cells express the parietal phenotype. However, if they remain in contact with undifferentiated EC cells (for example, in suspension aggregates, or in high-density monolayer cultures), they receive signals promoting expression of the visceral endoderm phenotype. According to this model, the first endoderm cells to differentiate from F9 EC cells are analogous to the primary endoderm cells of the 4.5-day mouse embryo, and modulation of their phenotype in culture may therefore provide clues about the nature of the intercellular interactions thought to be involved in endoderm differentiation during normal embryogenesis[14].

Electron microscopy was performed by Rita Tilly.

Received 12 December 1980; accepted 5 March 1981.

1. Sherman, M. I. & Miller, R. A. *Devl Biol.* **63,** 27–34 (1978).
2. Strickland, S. & Mahdavi, V. *Cell* **15,** 393–403 (1978).
3. Solter, D., Shevinsky, L., Knowles, B. B. & Strickland, S. *Devl Biol.* **70,** 515–521 (1979).
4. Strickland, S. & Sawey, M. J. *Devl Biol.* **78,** 76–85 (1980).
5. Jetten, A. M., Jetten, M. E. R. & Sherman, M. I. *Expl Cell Res.* **124,** 381–391 (1979).
6. Strickland, S., Smith, K. K. & Marotti, K. R. *Cell* **21,** 347–355 (1980).
7. Enders, A. C., Given, R. L. & Schlafke, S. *Anat. Rec.* **190,** 65–78 (1978).
8. Gardner, R. L. in *Results and Problems in Cell Differentiation* Vol. 9 (ed. Gehring, W. J.) 205–241 (Springer, Berlin, 1978).
9. Adamson, E. D. & Ayers, S. E. *Cell* **16,** 953–965 (1979).
10. Hogan, B. L. M., Cooper, A. & Kurkinen, M. *Devl Biol.* **80,** 289–300 (1980).
11. Howe, C. C. & Solter, D. *Devl Biol.* **77,** 480–487 (1980).
12. Hogan, B. L. M. *Devl Biol.* **76,** 275–285 (1980).
13. Dziadek, M. & Adamson, E. *J. Embryol. exp. Morph.* **43,** 298–313 (1978).
14. Hogan, B. L. M. & Tilly, R. *J. Embryol. exp. Morph.* (in the press).
15. Martin, G. R., Wiley, L. M. & Damjanov, I. *Devl Biol.* **61,** 230–244 (1977).
16. Lo, C. W. & Gilula, N. B. *Devl Biol.* **75,** 93–111 (1980).
17. Solter, D., Damjanov, I. & Skreb, N. *Anat. entwgesch. Monogr.* **132,** 291–298 (1970).
18. Laemmli, U. K. *Nature* **227,** 680–685 (1970).

Aromatic-dependent *Salmonella typhimurium* are non-virulent and effective as live vaccines

Susan K. Hoiseth & B. A. D. Stocker

Department of Medical Microbiology, Stanford University School of Medicine, Stanford, California 94305, USA

An auxotrophic mutant which requires a metabolite not available in vertebrate tissues should be unable to grow in such tissue and thus be non-virulent. Most mutations to auxotrophy do not affect virulence[1-3], presumably because the required metabolites are available at sufficient concentration. However, Bacon *et al.*[1,2] noted that a *Salmonella typhi* mutant requiring *p*-aminobenzoic acid (*p*AB) was less virulent (for mice) than its parent. *Salmonella*, unlike vertebrates, cannot assimilate exogenous folate and must synthesize it from *p*AB; the virtual absence of *p*AB from vertebrate tissues is shown by the efficacy of sulphonamide chemotherapy. Yancey *et al.*[4] reported reduced virulence for a *S. typhimurium* mutant with a requirement for 2,3-dihydroxybenzoate (DHB), the precursor of the bacterial iron-acquisition compound, enterochelin. As DHB is not a known vertebrate metabolite, it would be expected to be absent from mouse tissues. *Salmonella* synthesize *p*AB and DHB from chorismate, the final product of the aromatic biosynthetic (*aro*) pathway. A complete block at any step of this pathway should make *S. typhimurium* auxotrophic for two compounds not available in vertebrate tissues, and thus non-virulent. We describe here the use of a tetracycline-resistance transposon, Tn*10* (refs 5, 6), inserted in gene *aroA* to produce non-reverting, aromatic-requiring derivatives of virulent *S. typhimurium* strains. These derivatives were virtually non-virulent; their use as live vaccines conferred excellent protection against challenge with a virulent strain.

The *aro* transposon insertion used was *aroA554*::Tn*10*, in the chromosome of *S. typhimurium* strain LT2. As wild-type LT2 is only weakly pathogenic we transduced the *aroA554*::Tn*10* region into a virulent *S. typhimurium* strain, using phage P22 HT105/1 *int* (ref. 7) and selecting for tetracycline resistance. The virulent strain used as recipient, SL3201, is a mouse-passaged re-isolate of the genetically marked mouse-virulent *S. typhimurium* strain SL4522, of biotype FIRN[8]. Two of the transductants obtained, SL3217 and SL3218, required tryptophan, tyrosine and phenylalanine. Their additional requirements for *p*AB and DHB were observed on defined medium, made with purified agar to minimize *p*AB contamination, and containing citrate, a known iron chelator, to reduce iron availability; DHB was not needed if the medium was supplemented with ferrous sulphate. Strains SL3217 and SL3218 were then tested for pathogenicity by intraperitoneal (i.p.) injection of bacteria (grown in nutrient broth with tetracycline at 5 μg ml^{-1}) into adult mice of the *Salmonella*-susceptible[9] inbred line C57BL. Mice given 3×10^5 live bacteria of either of the two *aroA554*::Tn*10* strains showed no ill effects; the LD_{50} for the *aro*$^+$ parent strain, SL3201, was <20 bacteria.

Strains auxotrophic because of Tn*10* insertion in a biosynthetic gene can regain their lost biosynthetic ability and lose tetracycline resistance by 'clean excision' or deletion of the transposon[5,6]; for strains SL3217 and SL3218 the frequency of reversion to aromatic independence (and tetracycline sensitivity) was $\sim10^{-8}$ per bacterium per generation. Such *aro*$^+$ revertants had regained virulence ($LD_{50}<20$, i.p.). To obtain non-reverting *aro*$^-$ strains we made use of the tendency of Tn*10* to cause DNA alterations within the transposon itself, sometimes extending into adjacent chromosomal genes. Many such deletion or deletion–inversion events cause both loss of tetracycline resistance and inability of the affected gene to revert to the wild-type form[10,11]. Tetracycline-sensitive variants of the *aroA554*::Tn*10* transductant SL3218 were found amongst the survivors in a culture exposed to ampicillin at 2,500 μg ml^{-1} while growing in the presence of tetracycline (5 μg ml^{-1}), a concentration bacteriostatic for cells lacking Tn*10*. A tetracycline-sensitive variant thus obtained, SL3235, did not revert to *aro*$^+$ at detectable frequency ($<10^{-11}$ per bacterium per generation). We also derived an *aroA554*::Tn*10* transductant, SL1346, from a genetically marked subline, SL1344, of *S. typhimurium* strain S2337; this strain differs from SL3201 both in biotype and in that it is virulent for calves[12]. A non-reverting, tetracycline-sensitive variant, SL3261, was isolated from strain SL1346 by plating on nutrient agar containing autoclaved chlortetracycline and fusaric acid, a combination which inhibits growth of tetracycline-resistant bacteria[13]. Both the stable *aroA*$^-$ strains, SL3235 and SL3261, were non-virulent for mice of another *Salmonella*-susceptible inbred line, BALB/c (ref 9) (no deaths from inocula of $\sim3\times10^6$ bacteria i.p., compared with LD_{50} of <20 for the *aro*$^+$ parent strains).

In addition to the requirement for aromatic amino acids, *p*AB and DHB, *aro* mutants will lack two other products of chorismate: ubiquinone, made via *p*-hydroxybenzoate, and menaquinone, made via *o*-succinylbenzoate. Lack of some chorismate-derived metabolite may account for two recently noted consequences of *aro* defect: failure to hypermodify a uridyl base in tRNA, reported by Björk[14] for *aro* mutants in *Escherichia coli*, and failure to produce H_2S from thiosulphate, noted by us for *S. typhimurium* with non-leaky blocks at *aroA* or *aroD*. We considered the possibility that some such consequence of *aro* defect might account, in part or entirely, for the non-virulence of *aro*$^-$ strains of *S. typhimurium*. However, a preliminary experiment showed that a large inoculum (2×10^6 bacteria i.p.) of the non-virulent *aro*$^-$ strain, SL3261, caused fatal infections in most BALB/c mice provided with drinking water containing both *p*AB and DHB; the same inoculum caused no apparent ill effects in mice which received *p*AB alone or DHB alone, or neither benzoate. This suggests that requirement for *p*AB and for DHB each by itself suffices to make strain SL3261 non-virulent for mice maintained on a normal diet, and that inability to synthesize other products of chorismate is not relevant to the non-virulence of *aro*$^-$ *S. typhimurium*.

Vaccination by injection of killed *Salmonella* confers only poor protection (reduction in mortality) against challenge with a virulent strain[9]. We therefore tested the ability of our stable *aro*$^-$ strains, given i.p. as live vaccine, to protect against later challenge with a virulent *S. typhimurium* strain, given i.p. or by feeding (Table 1). The *S. typhimurium* strain used as challenge, SL1344 (the genetically marked virulent grandparent of *aro*$^-$

Table 1 Challenge by i.p. or oral administration of virulent *S. typhimurium* strain SL1344 (*aro*$^+$) to BALB/c mice immunized by i.p. injection of live *aro*$^-$ *S. typhimurium*

aro$^-$ live vaccine Strain	Dose	Challenge by *aro*$^+$ strain SL1344 Deaths/No. tested (days to death)
a Challenge by injection of 5×10^5 bacteria i.p. 34 days after live vaccine		
Controls	None	5/5 (day 2, 3, 3, 4, 5)
SL3261	2×10^4 i.p.	1/5 (day 3)*
	2×10^5 i.p.	0/5*
	2×10^6 i.p.	0/5*
b Challenge by feeding 3×10^7 bacteria 30 days after live vaccine		
Controls	None	5/5 (day 7, 8, 9, 9, 12)
SL3261	3×10^4 i.p.	0/5†
	3×10^5 i.p.	0/5†
	3×10^6 i.p.	0/5†
SL3235	3×10^4 i.p.	0/5†
	3×10^5 i.p.	0/5†
	3×10^6 i.p.	0/5†

Mice used were 15–17-week-old BALB/c males.
* All surviving mice were apparently well on day 60 after challenge.
† All mice alive and apparently well on day 120 after challenge.

0028-0836/81/210238—02$01.00

live vaccine strain SL3261), is virulent for mice, given i.p. or by feeding, and for calves by the oral route[12]. The mice used (the susceptible BALB/c line) were given graded i.p. inocula of one or the other stable aro^- strain and were challenged 1 month later with the virulent strain, either 5×10^5 bacteria i.p. or 3×10^7 bacteria given by feeding (on bread cubes, after a 12-h fast). Mice challenged by feeding were housed one per cage to prevent cross-infection. Each of the two aro^- strains, given i.p. in doses of 3×10^4, 3×10^5 or 3×10^6 (five mice per dose), gave complete protection against the oral challenge (no deaths and no apparent illness in 120 days of observation), whereas all five non-immunized control mice were dead by day 12. The two larger doses (2×10^5 and 2×10^6) of the aro^- strain SL3261 protected all mice (five mice per dose) against i.p. challenge and the smallest dose used, 2×10^4 live bacteria, protected four of five mice (Table 1). Thus i.p. injection of as few as 2×10^5 live aro^- bacteria into mice of a *Salmonella*-susceptible inbred line, BALB/c, gave complete protection against later i.p. injection of 5×10^5 virulent *S. typhimurium* (that is, >25,000 LD_{50}).

These results are in striking contrast to the failure of killed vaccines to protect such mice; Robson and Vas[9] found that six i.p. doses of killed *S. typhimurium* vaccine failed to protect BALB/c (or C57BL) mice against the lethal effect of i.p. injection of even as few as 10 cells of a virulent *S. typhimurium* strain. Oral administration of live bacteria of a non-virulent *galE* mutant of *S. typhi* has recently been shown to prevent typhoid fever in school children in Egypt[15]. Taken together, the data suggest that live vaccines may be useful for protecting man or domestic animals against *Salmonella* infections. Our results further suggest that strains with complete, non-reverting blocks in aromatic biosynthesis may be appropriate for such purposes.

This work was supported by US PHS grant AI07168 from the National Institute of Allergy and Infectious Diseases, by US Army Medical Research and Development Command contract DAMD-17-79-C-9076, and by a research contract with SmithKline Animal Health Products. S.K.H. was supported by National Research Service Award no. NIGMS 07276 and a NSF predoctoral fellowship. We thank Dr J. Roth of the University of Utah for the LT2 *aroA554*::Tn*10* insertion strain.

Received 19 November 1980; accepted 2 March 1981.

1. Bacon, G. A., Burrows, T. W. & Yates, M. *Br. J. exp. Path.* **31,** 714–724 (1950).
2. Bacon, G. A., Burrows, T. W. & Yates, M. *Br. J. exp. Path.* **32,** 85–96 (1951).
3. Herzberg, M. *J. infect. Dis.* **111,** 192–203 (1962).
4. Yancey, R. J., Breeding, S. A. & Lankford, C. E. *Infect. Immun.* **24,** 174–180 (1979).
5. Kleckner, N., Chan, R. K., Tye, B. & Botstein, D. *J. molec. Biol.* **97,** 561–575 (1975).
6. Kleckner, N. *Cell* **11,** 11–23 (1977).
7. Schmieger, H. *Molec. gen. Genet.* **119,** 75–88 (1972).
8. MacPhee, D. G. & Stocker, B. A. D. *J. Bact.* **100,** 240–246 (1969).
9. Robson, H. G. & Vas, S. I. *J. infect. Dis.* **126,** 378–386 (1972).
10. Kleckner, N., Roth, J. & Botstein, D. *J. molec. Biol.* **116,** 125–159 (1977).
11. Kleckner, N., Reichardt, K. & Botstein, D. *J. molec. Biol.* **127,** 89–115 (1979).
12. Wray, C., & Sojka, W. J. *Res. vet. Sci.* **25,** 139–143 (1978).
13. Bochner, B. R., Huang, H., Schieven, G. & Ames, B. N. *J. Bact.* **143,** 926–933 (1980).
14. Björk, G. R. *J. molec. Biol.* **140,** 391–410 (1980).
15. Wahdan, M. H. *et al. Bull. Wld Hlth Org.* **58,** 469–474 (1980).

Denervated endplates have a dual population of junctional acetylcholine receptors

Tami A. Levitt & Miriam M. Salpeter

Section of Neurobiology & Behavior, and School of Applied & Engineering Physics, Cornell University, 146 Clark Hall, Ithaca, New York 14853, USA

The acetylcholine receptors (AChRs) at the neuromuscular junctions of mature innervated muscle are distinguished, in part, from extrajunctional receptors of embryonic muscle fibres by their metabolic stability[1–5] and high concentration at the tops of the postjunctional folds[6,7]. After denervation, the rate of degradation of receptors at the neuromuscular junction progressively increases[8–12] even though their local concentration does not change[7,8,11]. Two alternatives suggest themselves—there may be a simultaneous, progressive increase in degradation rate of all junctional receptors after denervation, or the increase may reflect the composite behaviour of two populations of junctional receptors with different metabolic stabilities: those existing before denervation ('original AChRs') and those inserted afterwards ('new AChRs'). Our study has been designed to distinguish between these possibilities. We have now found that up to 8–10 days after denervation, the original receptors have a degradation half life of about 8 days which is indistinguishable from that of innervated junctional receptors. After that time, their degradation half life increases to about 2.5 days. However, all new AChRs, introduced at junctional sites after denervation, have a half life characteristic of extrajunctional receptors (about 1 day). Our data thus support the hypothesis of a composite population of receptors at denervated neuromuscular junctions.

The turnover rate of junctional AChRs (either denervated or innervated) was assessed by the loss of radioactivity bound to the neuromuscular junction after labelling with ^{125}I-α-bungarotoxin, the virtually irreversible inhibitor of the muscle's nicotinic AChR[13,14]. The validity of this procedure has been established in previous studies[1–3,15–19], although a slow unbinding of α-bungarotoxin has been reported in culture[3]. We first compared the degradation rate of the receptors at innervated junctions with that of the receptors at denervated junctions that were present before denervation (original AChRs). Adult mice (27–37 g) were injected with ^{125}I-α-bungarotoxin intravenously (tail vein), 2.5 μg per 100 g body weight (the LD_{50} was established to be about 3.5 μg per 100 g). At 3 h after the injection, the right sternomastoid muscle was denervated under Nembutal anaesthesia by removing 1–2 mm of nerve as close to the muscle as possible, leaving the contralateral muscle as an innervated control. (In control animals, the specific junctional counts in the left and right muscles are not significantly different at the time of injection.) At 1–2-day intervals up to 20 days later, three to nine animals were anaesthesized and killed by intracardial perfusion with 4% paraformaldehyde. The two sternomastoid muscles were removed and stained for acetylcholinesterase[20] to identify the endplate band. Each muscle was

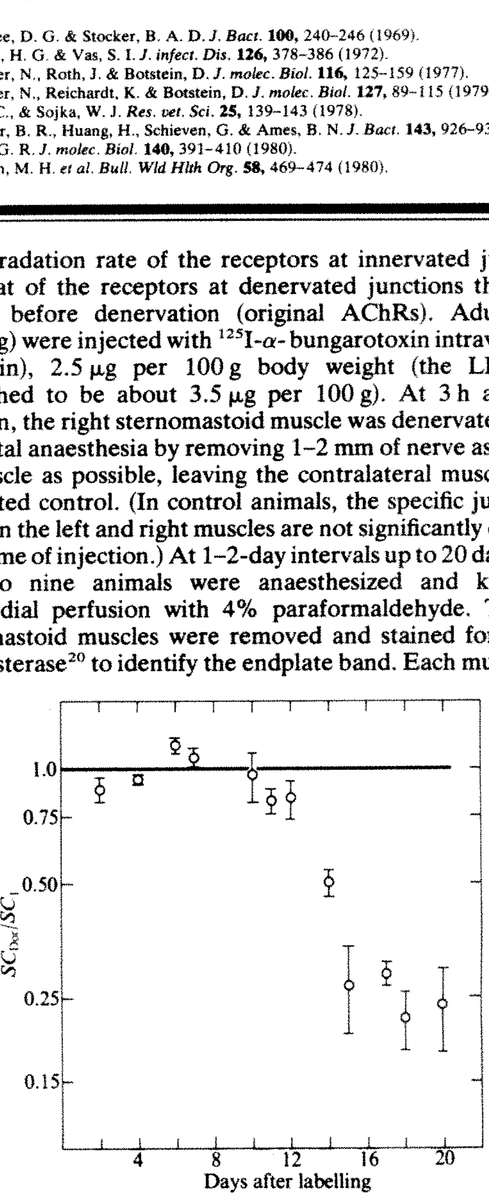


Fig. 1 Mice were injected with ^{125}I-α-bungarotoxin 3 h before denervation of the right sternomastoid muscle, thus labelling the original population of AChRs present before denervation. The ratio of specific endplate label of the denervated (SC_{Dor}) and innervated (SC_I) muscles was determined at different days thereafter. Values become significantly different from unity (solid line) after ~10 days of denervation. Error bars are s.e. of the mean of the ratios for three to nine animals per point.

0028-0836/81/210239—03$01.00

cut into three pieces, one containing the stained endplates, and the residual radioactivity of each piece was assayed by γ counting. The specific endplate radioactivity was determined by subtracting the radioactivity bound to the muscle pieces without endplates, from the activity in the piece containing the endplates on a per weight basis. We will call the specific endplate label at innervated junctions SC_I, that at denervated junctions which is due to the original receptors labelled before denervation SC_{Dor}, and that due to the total AChR population labelled any time after denervation SC_D.

Figure 1 plots the ratio of residual label, SC_{Dor}/SC_I, with time after denervation. This ratio is not significantly different from unity up to 8–10 days after denervation but becomes significantly lower thereafter. The early part of the post-denervation curve is consistent with the work of Berg and Hall[4], who observed that the decay rate of original AChRs does not change up to 5 days after denervation. The drop in the ratio (Fig. 1) indicates an increase in degradation rate of the original junctional AChRs by 10 days after denervation.

The degradation half lives of these innervated and denervated junctional receptors, labelled at the time of denervation, was determined as shown in Fig. 2. The label remaining with time at the innervated junctions (Fig. 2*a*) shows a single exponential decay (fit by linear regression) having a half time of 7.5 ± 1 days, similar to that previously reported[1–5,8]. The residual label at the denervated junctions (Fig. 2*b*) due to original receptors shows (by linear regression) two exponential decays, with a half life of 8.5 ± 3 days until ~9 days after denervation, and of 2.5 ± 0.5 days thereafter. (From our data, we cannot determine whether the degradation rate of the original receptors changes sharply at 9 days or whether, as may be more likely, there is a gradual shift from an 8-day to a faster degradation half time.)

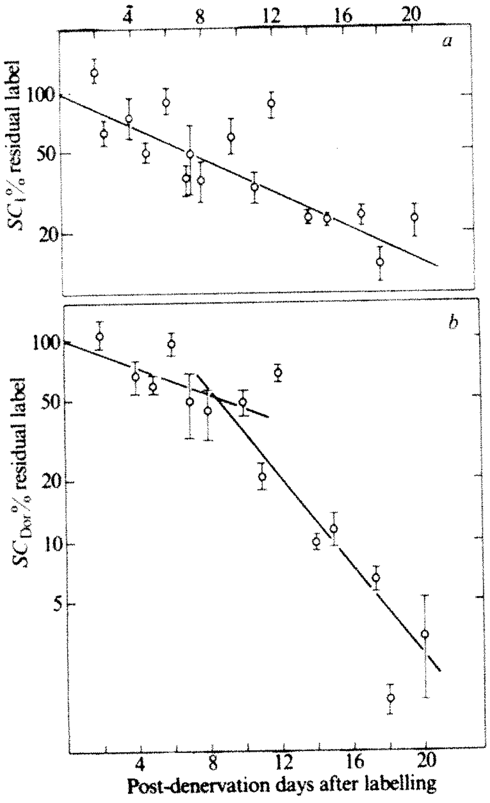


Fig. 2 Specific endplate label at different days after labelling (from all animals included in Fig. 1, plus a few additional innervated points) plotted separately for the innervated SC_I (*a*) and denervated SC_{Dor} (*b*) junctions. All curves are normalized to 100% at time of labelling, error bars are s.e. of the mean for three to nine animals per point. *a*, The innervated AChRs have a $t_{1/2}$ of 7.5 ± 1 days (by linear regression). *b*, The values for denervated original receptors can be fitted by two regression lines giving a degradation $t_{1/2}$ of 8.5 ± 3 days until 9 days after denervation and one of 2.5 ± 0.5 days thereafter. (As the $t_{1/2}$ of the original denervated receptors up to 9 days after denervation is not significantly different from that of innervated AChRs, an average value of 8 days will be used for both receptors during that period.) Although a gradual change from $t_{1/2}$ of about 8 days to a faster rate can also fit the data, we will, for convenience, assume that $t_{1/2}$ changes sharply and use the following mathematical formulation in subsequent calculations to assess the fraction of original receptors remaining at any given time (t) after denervation. When $t \leq 9$ days: $f(t) = 2 \exp -(t/8.0)$, and when $t > 9$ days: $f(t) = 2 \exp -[(9/8.0 + (t-9)/2.5)]$.

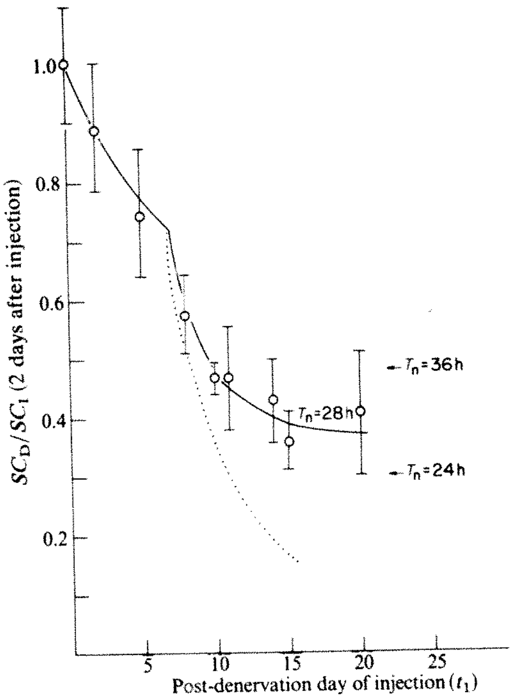


Fig. 3 Mice were injected with ^{125}I-α-bungarotoxin at different days (t_1) after denervation, thus labelling the AChRs as well as the remaining original population. The ratio of specific endplate label at denervated to innervated junctions (SC_D/SC_I) was determined 2 days after labelling. Experimental values (○) are from ref. 12 (with a few added points for confirmation). Curves were calculated on the basis of two alternative hypotheses: (1) (solid line), assuming that the junctional AChRs at denervated junctions consist of a dual population with different metabolic characteristics: the original receptors degraded as given in Fig. 2*b*, and new receptors, with a half life T_n. A value for T_n of 28 h gave the best fit to the data. The experimental data are within the range of values for T_n of 24–36 h (arrows) commonly associated with extrajunctional AChRs[1–5,15,16,18,19,21,22]. (2) (dotted line), assuming a dual population as given in (1) above, but with turnover rate of the original AChRs remaining with a $t_{1/2}$ of 8.0 days throughout, and the decreasing ratios in Fig. 1 representing a progressive decrease in the overall size of the denervated endplates. This model clearly does not fit the data. Additional evidence that the endplate does not decrease in overall number of sites was obtained (as reported in the text) from the ratio (SC_D/SC_I) of 1.3 ± 0.3 for 16-day denervates 3 h after labelling.

To test whether the drop in ratio SC_{Dor}/SC_I in Fig. 1 could be due to a decrease in the overall size of the denervated junctions as the denervated muscle atrophied, we injected four animals 16 days after denervation, waited 3 h to allow for removal of nonspecific label, removed the muscles and determined the specific endplate label by subtracting non-endplate label as described above. If there is no decrease in the number of AChRs at the denervated junction, the ratio (SC_D/SC_I) 3 h after injection should not be significantly different from unity. A value of 1.3 ± 0.3 (s.e.m.) was obtained. This is in agreement with Frank *et al.*[21], who found no change in toxin binding at denervated endplates for up to 2.4 weeks after denervation. As electron microscope autoradiography has shown that the AChR site density remains unchanged up to 16 days after denervation[8], we conclude that the total size of the denervated junction did not decrease. Additional evidence that the denervated endplate size did not decrease is presented in Fig. 3. Thus, the increased degradation rate of original AChRs by 8–10 days after denervation reflects an increased turnover rate (that is, as the original receptors are degraded, they are replaced by new AChRs on a one-to-one basis).

However, the behaviour of the original AChRs cannot fully explain the behaviour of total junctional AChRs at denervated junctions. We previously[12] examined the turnover of total junctional AChRs labelled at different times after denervation, and found that these total junctional AChRs begin to show a progressive increase in their turnover rate within 2 days of denervation, and reach a turnover half time of 30 h by 15 days. Even a week after denervation, when the original junctional receptors still decay slowly (with a $t_{1/2}$ of ~8 days, as given in Fig. 2*b*), the average turnover half life for total junctional receptors at the denervated junction had increased considerably (to ~2 days)[12]. A similar fast turnover rate was obtained for the total junctional AChRs specifically located at the postjunctional membrane of 8-day denervated neuromuscular junctions by electron microscope autoradiography[8], indicating that the faster turnover time reported for total junctional receptors is not an artefact of the γ-counting procedure.

To reconcile the data obtained for the original receptors in Fig. 2*b* with those for the total junctional receptors at denervated junctions reported previously[8,12] and in Fig. 3, we tested the assumption that after denervation, the original junctional receptors are degraded with the overall time course described in Fig. 2*b*, but that they are replaced by new receptors with a faster turnover half life (T_n). Based on this model we calculated the expected SC_D/SC_I for the experimental conditions of our earlier study[12]. In that earlier study, the average turnover rate of total junctional receptors at denervated neuromuscular junctions was obtained over successive 2-day periods by injecting animals with ^{125}I-α-bungarotoxin on different days after denervation (t_1) and determining the residual junctional label 2 days later (Δt). Therefore, in that study, the contribution to SC_D due to the original receptors labelled at time t_1 and evaluated at time Δt is $f(t_1)\times 2\ \exp-(\Delta t/t_{1/2})$. The $t_{1/2}$ is either 8 or 2.5 days, or a combination of both (depending on the value of t_1); and $f(t_1)$ is the fraction of original receptors remaining at the junction at time t_1 (Fig. 2*b*). In addition, as the absolute site density of AChRs remains constant after denervation[8], the new receptors available for labelling at time t_1 are $[1-f(t_1)]$, and their contribution to the specific endplate label (SC_D) at time $(t_1+\Delta t)$ is $[1-f(t_1)]\times 2\ \exp-(\Delta t/T_n)$, where T_n is the half life of the new receptors, and is the only free parameter in this model. Figure 3 plots the experimental points for SC_D/SC_I for successive 2-day periods (Δt) as obtained in our previous study (plus new data for confirmation) and compares it with the theoretical curve (solid line) based on the above model, using $T_n = 28$ h for best fit to the data. Figure 3 shows that this model reproduces the shape of the experimental results, with the T_n well within the range reported for extrajunctional AChRs[1-5,15,16,18,19,21,22]. Reiness *et al.*[3] have reported a slow unbinding of α-bungarotoxin from AChRs *in vitro*. If we assume a similar unbinding rate ($t_{1/2}$ of 13 days) in our system *in vivo*, and include this value in the above model, the shape of the theoretical curve does not fit the experimental data. Therefore, if unbinding of toxin is occurring in our system, it must be at a considerably slower rate than that seen by Reiness *et al.*[3]. Another model (dotted curve) which assumes a decrease in size of the denervated junction also does not fit the data.

We conclude that the denervated neuromuscular junction contains a dual population of junctional receptors with distinct metabolic characteristics. The first consists of junctional AChRs which are present at the time of denervation and start with a $t_{1/2}$ of about 8 days and eventually (by about 9 days) lose their metabolic stability. The second consists of new receptors, which have the same metabolic characteristics as extrajunctional receptors and replace (on a one-to-one basis) the original receptors as these are degraded. Earlier studies have already shown that the metabolic stability and high-density clustering of junctional AChRs are not linked either in development or after denervation[8-10,23,24]. The nerve seems to have a different role in determining these two properties of junctional AChRs. Initially it is required to induce the high-density AChR localization on the postjunctional membrane. Once established, however, this localization is maintained and can even be regenerated in the absence of the nerve[25-27]. The present study, on the other hand, shows that a continued presence of the nerve (or some nerve-related activity) is necessary for junctional receptors both to acquire and subsequently maintain metabolic stability, even in mature muscle.

This research was supported by grant NS09315 from the National Institute of Neurology and Communicative Disorders and Stroke, National Institutes of Health. We thank Ralph Loring for his help during this study. Presented in preliminary form at Neuroscience meeting, November 1980.

Received 12 November 1980; accepted 26 February 1981.

1. Chang, C. C. & Huang, M. C. *Nature* **253,** 643–644 (1975).
2. Heinemann, S., Merlie, J. & Lindstrom, J. *Nature* **274,** 65–68 (1978).
3. Reiness, C. G., Weinberg, C. B. & Hall, Z. W. *Nature* **274,** 68–70 (1978).
4. Berg, D. K. & Hall, Z. W. *J. Physiol., Lond.* **252,** 771–789 (1975).
5. Fambrough, D. M. *Physiol Rev.* **59,** 165–227 (1979).
6. Fertuck, H. C. & Salpeter, M. M. *J. Cell Biol.* **69,** 144–158 (1976).
7. Porter, C. W. & Barnard, E. A. *Expl Neurol.* **48,** 542–550 (1975).
8. Loring, R. & Salpeter, M. M. *Proc. natn. Acad. Sci. U.S.A.* **77,** 2293–2297 (1980).
9. Brett, T. S. & Younkin, S. G. *Neurosci. Abstr.* **5,** 2568 (1979).
10. Loring, R. H. & Salpeter, M. M. *J. Cell Biol.* **83,** 137a (1979).
11. Loring, R. H. & Salpeter, M. M. *Neurosci. Abstr.* **4,** 1930 (1978).
12. Levitt, T. A., Loring, R. H. & Salpeter, M. M. *Science* **210,** 4469 (1980).
13. Chang, C. C. & Lee, C. Y. *Archs int. Pharmacodyn. Thér.* **144,** 241–257 (1963).
14. Lee, C. Y., Tseng, L. F. & Chiu, T. H. *Nature* **215,** 1177–1178 (1967).
15. Linden, C. D. & Fambrough, D. M. *Neurosci. Abstr.* **4,** 527–538 (1978).
16. Gardner, J. M. & Fambrough, D. M. *Cell* **16,** 661–674 (1979).
17. Fertuck, H. C., Woodward, W. & Salpeter, M. M. *J. Cell Biol.* **66,** 209–213 (1975).
18. Merlie, J. P., Changeux, J. P. & Gros, F. *Nature* **264,** 74–76 (1976).
19. Devreotes, P. N. & Fambrough, D. M. *J. Cell Biol.* **65,** 335–358 (1975).
20. Karnovsky, M. J. & Roots, L. *J. Histochem. Cytochem.* **12,** 219–221 (1964).
21. Frank, E.,Gautvik, K. & Sommerschild, H. *Acta physiol. scand.* **95,** 66–76 (1975).
22. Steinbach, J. H., Merlie, J., Heinemann, S. & Bloch, R. *Proc. natn. Acad. Sci. U.S.A.* **76,** 3547–3551 (1979).
23. Burden, S. *Devl Biol.* **61,** 79–85 (1977).
24. Burden, S. *Devl Biol.* **57,** 317–329 (1977).
25. Burden, S. J., Sargent, P. B. & McMahan, U. J. *J. Cell Biol.* **82,** 412–425 (1979).
26. Marshall, L. M., Sanes, J. R. & McMahan, U. J. *Proc. natn. Acad. Sci. U.S.A.* **74,** 3073–3077 (1977).
27. McMahan, U. J., Sanes, J. R. & Marshall, L. M. *Nature* **271,** 172–174 (1978).

Sexual dimorphism in a photoreceptor

N. Franceschini*†, R. Hardie*, W. Ribi*
& K. Kirschfeld*

* Max-Planck-Institut für biologische Kybernetik, 7400 Tübingen, FRG
† Institut de Neurophysiologie et Psychophysiologie, CNRS, 13277 Marseille Cedex 9, France

The recent observation of fluorescence emission from photoreceptor cells in flies[1] offers a new opportunity to map the distribution of the various spectral types in a highly organized retinal mosaic. Previous studies had led to the conclusion that in each ommatidium the two central receptor cells are quite distinct from their six neighbours in terms of anatomical projection[2,3], visual pigment and physiology[4]. We show here that in a restricted, well defined region of the male eye only, one of these central cells resembles its six neighbours not only in having the same visual pigment and physiological properties, but also in sending its axon to the same neuropile. A clue to the function of this sexually dimorphic retinal organization is given by the fact that these male-specific cells dominate the frontal-dorsal part of the retina (fovea) which is used by the male to track females in aerobatic chases[5,6]. We suggest that the male fovea, already known to drive sex-specific higher-order neurones[7], has sacrificed colour vision for other more vital kinds of information processing such as improving quantum catch and possibly also contrast transfer or movement detection.

Each ommatidium in the compound eye of the housefly (*Musca domestica*) contains eight photoreceptors which are

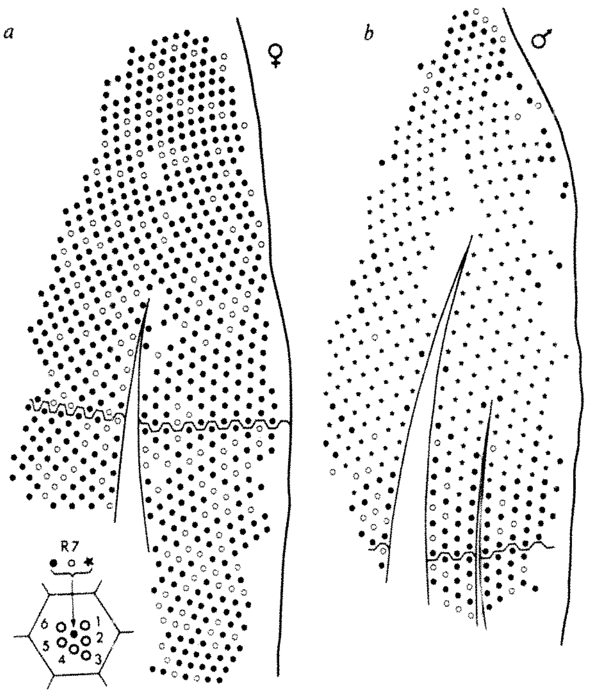


Fig. 1 Inset, arrangement and numbering of the rhabdomeres in a fly ommatidium (distal cross-section of a ventral ommatidium). Rhabdomere R8 is a proximal extension of R7. The inset indicates in brackets the three symbols used for R7 in the adjoining maps. *a*, *b*, Maps of the frontal part of the right retina showing the distribution of the three types of R7 receptor cells in the compound eye of a female (*a*) and male (*b*) housefly (*M. domestica*, white eye). Each dot or star represents the distal tip of rhabdomere R7 (see inset) as seen *in vivo* in each ommatidium, using epifluorescence microscopy (blue excitation) combined with optical neutralization of the cornea[1] (water immersion objective ×25/0, 65). Empty spaces represent ommatidia in which the fluorescence colour of R7 could not be clearly distinguished. ●, Green-fluorescing R7s (UV + blue-sensitive cells[4]); ○, non-fluorescing R7s (pure UV-sensitive cells[4]); asterisks, red-fluorescing R7s (dual peak, UV + blue-green-sensitive cells, see Fig. 2). The zig-zag line represents the equator of the eye, that is, a (virtual) line of dorso-ventral symmetry, about which the R1–6 rhabdomeric pattern (see inset) is inverted. About 70% green-fluorescing and 30% non-fluorescing R7s make up the female retina (*a*) and the ventral part of the male retina (*b*). The red-fluorescing R7s are present in the frontal-dorsal region of the male eye exclusively (*b*). Note how they gradually invade the retina above the equator. These maps (each of which encompasses approximately one-fifth of the fly retina) were reconstructed from 35(a) and 38(b) colour photographs with overlapping borders. Although such maps are most conveniently made in white-eyed flies, the situation seems to be qualitatively similar in wild type.

usually classified as six large peripheral cells, R1–6 surrounding two smaller central cells, R7 and R8 (see Fig. 1, inset). The photoreceptive segments (rhabdomeres) of the latter are arranged one above the other to form a single light-guiding structure[3]. As recognized by Cajal[2], the axons of R1–6 cells project only as far as the first optic ganglion (the lamina) whereas the axons of both R7 and R8 pass through this neuropile and eventually synapse first in the second optic ganglion (the medulla). A similar retinal dichotomy of long and short visual fibres seems to be a general feature of diurnal compound eyes.

The lamina is organized into optic cartridges each of which normally receives axons from six R1–6 cells 'looking' in the same direction of space, but belonging to different ommatidia[8,9]. The signal summation which occurs within each cartridge[10] necessarily increases the signal-to-noise ratio, and hence the sensitivity of these visual channels whose information is relayed to the medulla by second-order monopolar neurones[2,3,9,11]. As the individual properties of the single receptors are buried in this summation process, it is not surprising that the six cells feeding each cartridge all have the same (dual-peak) spectral sensitivity[12,13].

In contrast to the homogeneity of the R1–6 population throughout the retina, the population of central cells, R7 and R8, is highly heterogeneous. Three classes of receptor cells R7 have been found using *in vivo* epifluorescence microscopy. While R1–6 cells always fluoresce red under blue excitation, R7 rhabdomeres exhibit three distinct colours: red, green and black (that is, nonfluorescing), depending on the ommatidium[1].

We have screened large regions of the retina by rotating the fly stepwise on a precision goniometer under the immersion objective of the epifluorescence microscope, and constructed precise maps of the distribution of the R7 types from photographic series (Fig. 1). The most striking conclusion from these maps is that females do not possess any red-fluorescing R7 and that these are exclusively encountered in the frontal-dorsal part of the male eye where they account for between 85 and 95% of all R7s (asterisks in Fig. 1*b*).

This sex-specific retinal architecture stimulated a closer investigation of the properties of the red-fluorescing R7s. As their fluorescence colour was similar to that of R1–6 we thought they would have the same photopigment, and hence spectral sensitivity, as their six neighbours. Intracellular recordings and stainings revealed that the spectral sensitivities were almost indistinguishable (Fig. 2), a finding corroborated by microspectrophotometry of individual R7 rhabdomeres (ref. 14). Taken together, these two methods suggest that the red-fluorescing R7s have inherited not only the R1–6 rhodopsin, responsible for the peak near the 480 nm, but also apparently the UV-sensitizing pigment[13], responsible for the second peak near 360 nm (Fig. 2). The two other types of R7 have spectral sensitivities quite different from those of R1–6 cells[4].

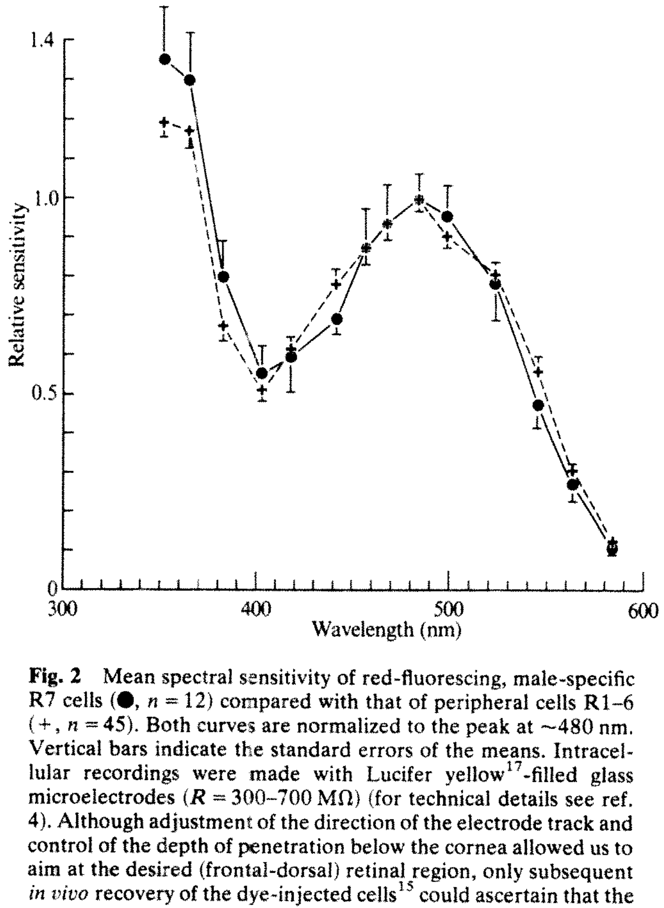


Fig. 2 Mean spectral sensitivity of red-fluorescing, male-specific R7 cells (●, $n = 12$) compared with that of peripheral cells R1–6 (+, $n = 45$). Both curves are normalized to the peak at ~480 nm. Vertical bars indicate the standard errors of the means. Intracellular recordings were made with Lucifer yellow[17]-filled glass microelectrodes ($R = 300$–700 MΩ) (for technical details see ref. 4). Although adjustment of the direction of the electrode track and control of the depth of penetration below the cornea allowed us to aim at the desired (frontal-dorsal) retinal region, only subsequent *in vivo* recovery of the dye-injected cells[15] could ascertain that the R7s recorded were indeed from a retinal region dominated by red-fluorescing R7s (asterisks in Fig. 1*b*). In contrast to the two other types of R7 of the retina[4] (Fig. 1), the male-specific R7 cells seem to have a spectral sensitivity similar to that of R1–6 cells. *M. domestica*, white eye.

Electrophysiologically the responses of the male-specific R7s are so similar to those of R1–6 cells that it proved impossible to tell them apart during recording and only *in vivo* recovery of blindly injected cells[15] revealed that we had recorded from them.

These R7 cells resemble R1–6 cells even at the ultrastructural level. Thus, electron micrographs showed that, compared with other R7 cells, the red-fluorescing ones have larger cell bodies and axons, and also wider rhabdomeres (Fig. 3)—all features which approximate to the situation in R1–6 cells. Most surprisingly, electron microscope serial sectioning established that, like R1–6 cells, the axons of these male-specific R7 cells also form connections with second-order neurones in the lamina (Fig. 3*c*), while axon profiles from stained cells revealed that their axons also stop short in the lamina (Fig. 3*g*) instead of projecting to the medulla like the other R7s of the eye.

In a restricted, well defined region of the eye of male houseflies, the R7 photoreceptors thus seem to have been programmed to resemble the peripheral R1–6 photoreceptors in terms of visual pigment, electrophysiological properties and many fine details of anatomy. Curiously, the underlying R8 photoreceptors in this region go only halfway in this respect. They apparently contain the same visual pigment as R1–6 but remain anatomically distinct, still retaining the long visual fibre input to the medulla[14].

In looking for a function for this intriguing sexual dimorphism we are attracted by the recently discovered sex-specific chasing behaviour of male flies[5,6,16] and the sex-specific neurones found in the lobula of male flies[7]. Both the behaviour and the neurones receive their main visual input from the region of the eye (the fovea) where the red-fluorescing R7s are located. The male's chasing behaviour involves the tracking of a small, fast-moving spot (the female fly) and, in visual tasks requiring high temporal resolution, light quanta are as vital as at low light levels[16]. In this respect the participation of a seventh, spatially and spectrally matched input to the summation process at the first synaptic level could be considered an adaptation to increase further the signal-to-noise ratio (although by a small factor) in a quantum-limited task of obvious biological importance.

However, additional functions are suggested by details of the lamina terminals of the male-specific R7s, revealed by intracellular staining. Their axons do not penetrate as deeply into the cartridge layer as those of R1–6. Furthermore, just before they terminate, they produce a right-angled branch which runs ~10 μm (that is, to about the next cartridge)—always towards the frontal margin of the eye (Fig. 3*g*). If this lateral process mediated inter-cartridge interactions, it could subserve a peripheral form of movement detection. Alternatively, it could modify the contrast transfer properties of second-order neurones by providing lateral inhibition.

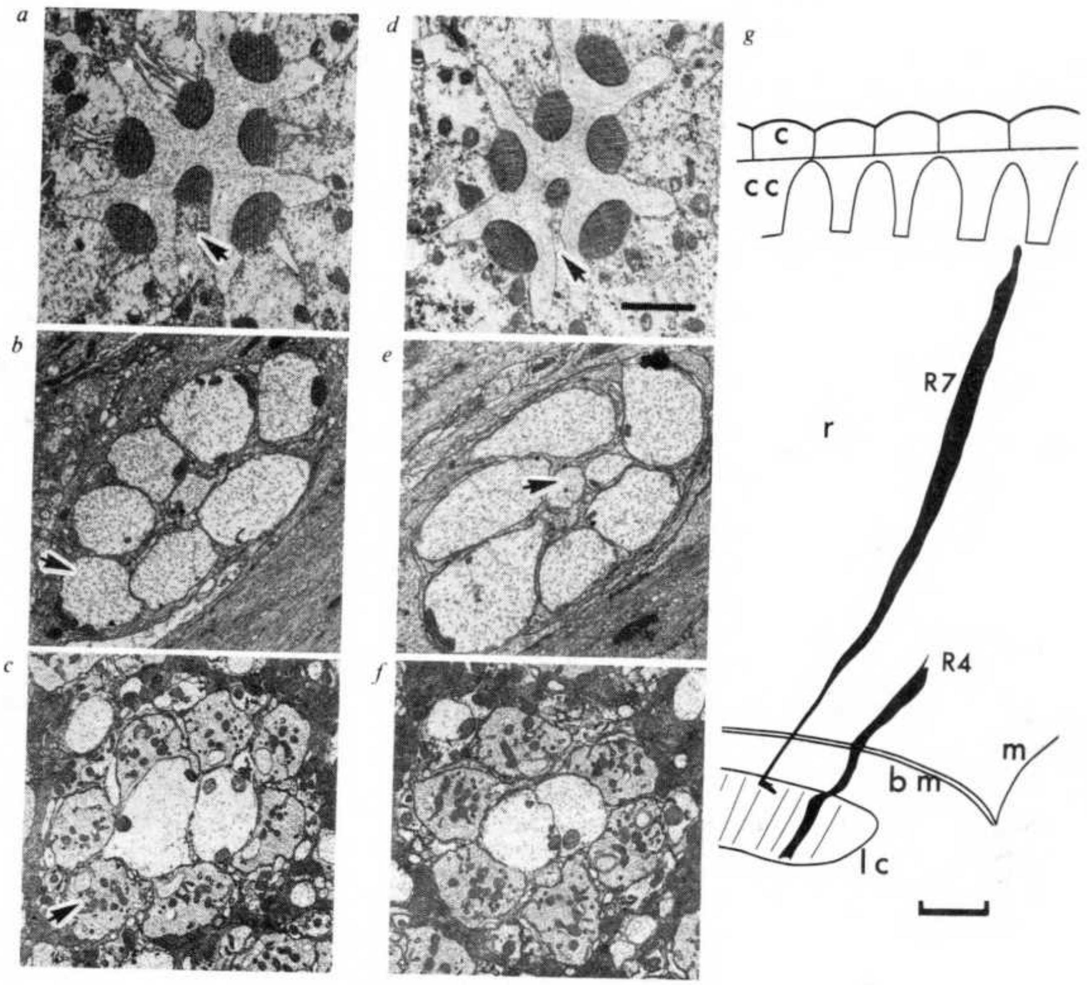


Fig. 3 *a–f*, Electron-micrographs of serial cross-sections which follow the receptors from two ommatidia: one containing a red-fluorescing R7 (*a–c*) and the other a green-fluorescing R7 (*d–f*). At the electron micrograph level, ommatidia containing non-fluorescing R7s (not shown) were indistinguishable from those containing green-fluorescing R7s. Before sectioning, the ommatidia were first identified from their *in vivo* fluorescence. Reliable correspondence between the fluorescence map and the electron micrograph map was achieved by the use of two landmarks recognizable by both methods—the equatorial line and a characteristically mis-oriented rhabdomere pattern. In each electron micrograph the R7 cell is indicated by an arrow. *a*, *d*, Cross-sections through the retina. Note that the rhabdomere and cell body of the red-fluorescing R7 (*a*) are larger than in the green-fluorescing R7 (*d*). *b*, *e*, Cross-sections just below the basement membrane showing the receptor axons as they leave the retina. Whereas the axon of the green-fluorescing R7 is small and surrounded by a ring of R1–6 axons (*e*), that of the red-fluorescing R7 is large and is itself a member of the ring (*b*). *c*, *f*, Cross-sections of the cartridges on the lamina. At this level the green-fluorescing R7 axon (not visible) has separated from the R1–6 axons which, alone, form a ring around two second-order monopolar cells (*f*). In contrast, the axon of the red-fluorescing R7 is a member of the ring and also forms connections with the second-order cells (*c*). *g*, Tracing of an entire male-specific, red-fluorescing R7 stained intracellularly with Lucifer yellow[17]. Note the shallow terminal in the lamina and the lateral process which is bent at right angles to the cell axis and runs ~10 μm before stopping. For comparison the terminal (only) of an R1–6 cell (actually R4) is shown, with the axon running, typically, to the bottom of the cartridge layer. c, Corneal facets; cc, crystalline cone; r, retina; m, frontal margin; bm, basement membrane; lc, lamina cartridge layer. Scale bar, 1 μm (*a–f*); 30 μm (*g*).

Whereas other parts of the retina possess an impressive array of colour receptors[4], the male fovea is necessarily colour blind. We suggest that this sacrifice is made to enhance one or more of the above mentioned processes (quantum catch, contrast transfer and movement detection), all of which can be seen as advantageous for the exacting task in which the male fovea seems to be mainly involved.

We thank Ms A. Rudolph and A. Totin for technical assistance and Drs M. M. Bate and D. G. Stavenga for critical reading of the manuscript. The dye Lucifer yellow was provided by Dr W. Stewart[17]. This work was supported by the MPG (FRG) and grants from the CNRS and DGRST (France).

Received 11 November 1980; accepted 19 February 1981.

1. Franceschini, N., Kirschfeld, K. & Minke, B. *Science* (in the press).
2. Cajal, S. R. *Trab. Lab. Invest. biol. Univ. Madr.* **7,** 217–256 (1909).
3. Trujillo-Cenoz, O. & Melamed, J. in *The Functional Organization of the Compound Eye* (ed. Bernhard, C. G.) 339–361 (Pergamon, London, 1966).
4. Hardie, R. C., Franceschini, N. & McIntyre, P. D. *J. comp. Physiol.* **133,** 23–39 (1979).
5. Land, M. F. & Collett, T. S. *J. comp. Physiol.* **89,** 331–357 (1974).
6. Wehrhahn, C. *Biol. Cybern.* **32,** 239 (1979).
7. Hausen, K. & Strausfeld, N. *Proc. R. Soc.* **B208,** 57–71 (1980).
8. Kirschfeld, K. *Expl Brain Res.* **3,** 248–270 (1967).
9. Braitenberg, V. *Expl Brain Res.* **3,** 271–298 (1967).
10. Scholes, J. *Kybernetik* **6,** 149–162 (1969).
11. Boschek, B. *Z. Zellforsch. mikrosk. Anat.* **118,** 369–409 (1971).
12. McCann, G. D. & Arnett, D. W. *J. gen. Physiol.* **59,** 534–558 (1972).
13. Kirschfeld, K., Franceschini, N. & Minke, B. *Nature* **269,** 386–390 (1977).
14. Hardie, R., Franceschini, N., Ribi, W. & Kirschfeld, K. (in preparation).
15. Franceschini, N. & Hardie, R. C. *J. Physiol., Lond.* **301,** 56P (1980).
16. Kirschfeld, K. & Wenk, P. *Z. Naturf.* **31,** 764–765 (1976).
17. Stewart, W. W. *Cell* **14,** 741–759 (1978).

Direct transfer of a lysosomal enzyme from lymphoid cells to deficient fibroblasts

I. Olsen, M. F. Dean, G. Harris & H. Muir

Kennedy Institute of Rheumatology, 6 Bute Gardens,
London W6 7DW, UK

Fibroblasts grown *in vitro* can secrete lysosomal enzymes into the culture medium which can then be taken up by endocytosis into other fibroblasts[1–4]. Such cells have been used as a source of enzyme for replacement therapy in the mucopolysaccharidoses[5–7], a group of related storage diseases resulting from inherited deficiencies in specific lysosomal enzymes, although the mechanism of action is unclear. Because lymphoid cells can interact directly with fibroblasts[8,9] and transfer DNA to them[10], we examined the possibility that a lysosomal enzyme could be transferred from normal lymphocytes to deficient fibroblasts by a mechanism other than secretion followed by endocytosis. We report here that fibroblasts from a patient with mucopolysaccharidosis type VII acquired β-glucuronidase (EC 3.2.1.31) activity as a result of direct cell-to-cell contact with lymphocytes.

Monolayer cultures of β-glucuronidase-deficient fibroblasts (GM 151 cells) contained no detectable enzyme activity, even at high cell density. After co-culture with lymphocytes, β-glucuronidase was detected in all the GM 151 cells examined (Fig. 1*a*). Direct measurement of enzyme activity acquired by the fibroblasts after interaction was not possible, because between 1 and 10% of the added lymphocytes adhered firmly to them even after repeated washing (Fig. 1*b*). Quantitative cytochemical assay in single GM 151 cells was possible, however, because the lymphocyte–fibroblast association was readily disrupted by trypsin.

The β-glucuronidase activity acquired by GM 151 cells increased rapidly during the first 24 h of co-culture, and decreased over the next several days (Fig. 2). In eight other separate experiments, maximum levels of between 29.9 and 48.1 nmol per h per 10^6 cells (mean 40.7 ± 2.1 s.e.) were acquired after 24 h.

The cytochemical assay accurately measured β-glucuronidase activity, as shown in Fig. 3. Note that the activity acquired by GM 151 cells (40.7 nmol per h per 10^6 cells) was above that in fibroblasts from two (clinically normal) people heterozygous for the β-glucuronidase allele, GM 1850 and GM 2074 (27.1 and 12.2 nmol per h per 10^6 cells, respectively), and also greater than that required for full correction of the metabolic deficiency in mucopolysaccharidosis type VII (ref. 11).

The β-glucuronidase activity acquired by 10^5 GM 151 cells was directly proportional to the number of lymphocytes added during 24 h of co-culture, up to 10^7 lymphocytes per culture dish (see Table 1). This optimum number of cells was therefore used for subsequent co-culture.

The persistence of acquired β-glucuronidase activity was examined. If, after the initial 24 h of co-culture, the non-adherent lymphocytes were replaced by fresh medium for a further 24 h, the β-glucuronidase activity decreased from an initial 40.7 nmol per h per 10^6 cells to 25.8, the same as that in fibroblasts incubated in the continued presence of lymphocytes (Fig. 2). Furthermore, the addition of 10^7 fresh lymphocytes after the initial 24 h of co-culture produced only a small further increase, to 33.1 nmol per h per 10^6 cells. The β-glucuronidase activity acquired by GM 151 cells during an initial 24 h of co-culture decreased even further during the next 48 h, either in the continued presence of the lymphocytes (19.4 nmol per h per 10^6 cells) or after they had been replaced by fresh medium (14.8 nmol per h per 10^6 cells). However, when the non-adherent lymphocytes were removed after 24 h of co-culture, and the GM 151 cells re-cultured for 48 h at high cell density, β-glucuronidase activity persisted at a very high level (39.4 nmol per h per 10^6 cells). This suggests that the decrease in activity after 24 h of co-culture was due to cell division rather than enzyme degradation.

Table 1 The effect of lymphocyte concentration on the appearance of β-glucuronidase activity after 24 h of co-culture with 10^5 GM 151 cells

Lymphocytes added per culture dish	β-Glucuronidase activity in GM 151 cells (nmol per h per 10^6 cells)
0	0
3×10^5	0
1×10^6	10.1
3×10^6	15.7
1×10^7	48.1
3×10^7	38.9

Monolayers of GM 151 cells were grown directly on 35-mm culture dishes and the co-culture initiated by the addition of the washed lymphocytes resuspended in 1 ml of Eagle's minimal essential medium (MEM), supplemented with fetal calf serum (FCS). After 24 h of co-culture, the non-adherent lymphocytes were removed and the adherent cells detached with 0.25% trypsin as described in Fig. 2 legend. After incubation with naphthol AS-BI-β-glucuronide and post-coupling with fast blue RR, quantitative cytochemistry of individual GM 151 cells was performed. The arithmetic means were calculated and converted to β-glucuronidase activity expressed as nmol per h per 10^6 GM 151 cells, as shown in Fig. 3.

It is not clear why prolonged co-culture or fresh lymphocytes did not increase further the β-glucuronidase activity acquired by GM 151 cells. Direct assay using 4-methylumbelliferyl β-D-glucuronide (see Fig. 3 legend), showed that the 10^7 lymphocytes added initially to 10^5 GM 151 cells, contained 10–40 nmol per h of β-glucuronidase activity (1–4 nmol per h per 10^6 lymphoid cells. After 24 h of interaction, almost all the lymphocytes were recovered, while the 10^5 GM 151 cells had acquired ~4 nmol per h of enzyme. This level of acquired enzyme, 40.7 nmol per h per 10^6 cells, was over 10 times above that of the lymphocytes, and was equivalent to 10–40% of the total lymphocyte activity initially added. However, preliminary experiments showed that there was also significant increase in the β-glucuronidase activity of the recovered lymphocytes themselves as a result of direct interaction with GM 151 cells. Furthermore, lymphocytes that had already been co-cultured could induce another round of enzyme acquisition when

0028-0836/81/210244—04$01.00

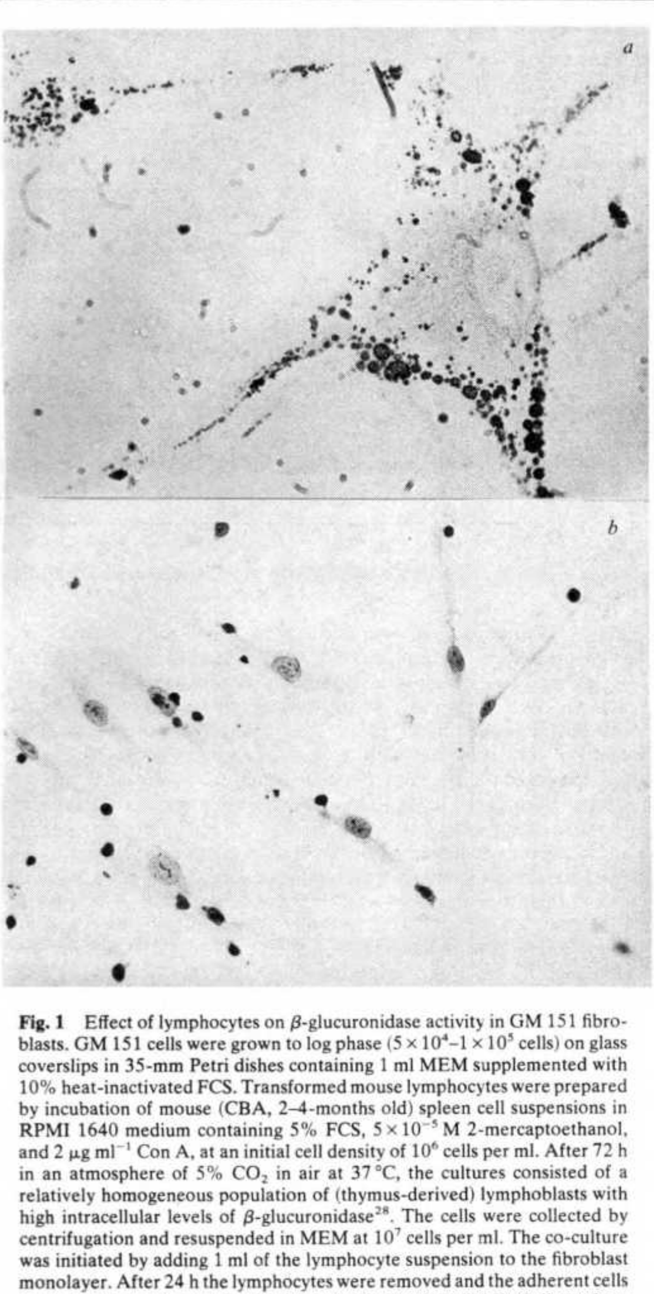


Fig. 1 Effect of lymphocytes on β-glucuronidase activity in GM 151 fibroblasts. GM 151 cells were grown to log phase (5×10^4–1×10^5 cells) on glass coverslips in 35-mm Petri dishes containing 1 ml MEM supplemented with 10% heat-inactivated FCS. Transformed mouse lymphocytes were prepared by incubation of mouse (CBA, 2–4-months old) spleen cell suspensions in RPMI 1640 medium containing 5% FCS, 5×10^{-5} M 2-mercaptoethanol, and 2 μg ml^{-1} Con A, at an initial cell density of 10^6 cells per ml. After 72 h in an atmosphere of 5% CO_2 in air at 37 °C, the cultures consisted of a relatively homogeneous population of (thymus-derived) lymphoblasts with high intracellular levels of β-glucuronidase[28]. The cells were collected by centrifugation and resuspended in MEM at 10^7 cells per ml. The co-culture was initiated by adding 1 ml of the lymphocyte suspension to the fibroblast monolayer. After 24 h the lymphocytes were removed and the adherent cells washed twice with 1 ml of phosphate-buffered saline (PBS). Cell numbers were determined by direct counting, using a haemocytometer. *a*, β-glucuronidase activity was measured as follows. Adherent fibroblasts were incubated with 1 ml of substrate (100 μg ml^{-1} naphthol AS-BI-β-D-glucuronide in 0.1 M acetate buffer, *p*H 4.5). After 3 h at 37 °C, the dishes were washed with PBS and 1.0 ml of freshly prepared fast blue RR (1 mg ml^{-1} in 0.01 M KP_i, *p*H 7.4) was added for 10 min at 4 °C. The dishes were washed with PBS, methanol, then distilled water and air dried. The coverslips were removed, fixed on to slides and mounted with Polymount for visual examination and photography. GM 151 cells cultured alone did not stain in these conditions. The accumulation of enzyme is seen in densely staining vesicular organelles which look like lysosomes. ×500. *b*, Adherent cells were fixed in methanol, washed with distilled water and air dried. The coverslips were removed, fixed on to slides and stained with Giemsa. The lymphocytes are smaller and more densely stained. Occasionally lymphoid cells were seen within GM 151 cells, but there was no sign of significant phagocytosis. ×175.

subsequently co-cultured for 24 h with a second monolayer of GM 151 cells. Thus the β-glucuronidase activity of the lymphoblast population was not exhausted by interaction, and the fibroblasts themselves seemed to limit acquisition of activity.

Quantities of β-glucuronidase secreted by cultured cells were determined by direct assay of the supernatant media using 4-methylumbelliferyl β-D-glucuronide as a substrate. Lymphocyte culture medium and medium from lymphocyte–GM 151 co-cultures both contained less than 5 nmol per h per ml of β-glucuronidase after 24 h. Incubation of GM 151 cells with these media for 24 h did not result in any detectable intracellular enzyme activity. The presence of a precursor to β-glucuronidase, which becomes enzymatically active after uptake by GM 151 cells, is therefore unlikely. This is consistent with the absence of such a precursor for this enzyme in murine macrophages[12]. The addition to GM 151 cells of partially purified β-glucuronidase from mouse 3T3 fibroblasts (S. Diment and M. F. Dean, unpublished data) also at 5 nmol per h per ml, was similarly without effect. Moreover, after the addition of cell-free enzyme extracts derived from 10^7 lymphocytes, GM 151 cells acquired barely detectable activity.

Endocytosis of released enzyme was further examined using mannose-6-phosphate, a competitive inhibitor of receptor-mediated uptake of several lysosomal hydrolases, including β-glucuronidase[13–15]. At 1 mM, it markedly inhibited endocytosis of exogenous enzyme prepared from 3T3 cells, but had no effect on the acquisition of β-glucuronidase by direct interaction with lymphocytes (Table 2). To produce comparable levels of enzyme activity in GM 151 cells, it was necessary to add much more exogenous β-glucuronidase derived from 3T3 cells than that present in 10^7 lymphocytes.

Concanavalin A(Con A)-mediated uptake[16] of enzyme secreted by the lymphocytes also cannot account for the acquisition of β-glucuronidase activity in GM 151 cells during co-culture, because the addition of Con A (2 μg ml^{-1}), either by itself, with the culture media, or with enzyme prepared from lymphocytes, had no promoting effect on β-glucuronidase activity in GM 151 cells. The effect of α-methylmannoside (α-MM) was next examined as it inhibits Con A-mediated phagocytosis of lysosomal enzymes[17]. Con A-transformed lymphocytes were washed in 100 mM α-MM and then co-cultured with GM 151 cells for 24 h in the presence of 25 mM α-MM. The level of β-glucuronidase acquired by GM 151 cells was almost identical to that of control cultures. Furthermore, co-culture with lymphocytes transformed by bacterial lipopolysaccharide similarly resulted in the appearance of enzyme activity in GM 151 cells. High levels of β-glucuronidase were also acquired when fibroblasts were co-cultured with cells not exposed to mitogens, such as human peripheral blood lymphocyte/monocyte preparations and monocyte-depleted mouse spleen cells.

Cells from patients with mucopolysaccharidosis type VII produce an inactive protein which cross-reacts with antibody prepared against active enzyme[18,19]. It is unlikely, therefore, that any factor secreted by the lymphocytes[20–23] led to the active expression of a previously repressed structural gene.

These experiments suggest that enzyme acquisition by GM 151 cells did not involve secretion of β-glucuronidase by lymphocytes and its subsequent uptake by GM 151 cells. To test

Table 2 β-glucuronidase activity in GM 151 cells: comparison of direct cell-to-cell contact with endocytosis of exogenous enzyme

	β-glucuronidase activity in GM 151 cells (nmol per h per 10^6 cells)		
Additions to 10^5 GM 151 cells	Direct contact	Millipore* chambers	Mannose-6-P (1 mM)
10^7 lymphocytes (15)	29.9	0	39.1
Exogenous 3T3 enzyme (200)	49.9	24.1	6.1

GM 151 cells were grown to log phase (1×10^5 cells) in 35-mm culture dishes; 1 ml of medium containing 10^7 lymphocytes or partially purified enzyme was added, and the cultures continued for 24 h. Values in parentheses indicate the β-glucuronidase activity added (total nmol per h per culture dish) in the form of intact lymphocytes or purified enzyme. These values were determined by direct biochemical assay using 4-methylumbelliferyl β-D-glucuronide substrate. β-glucuronidase activity was measured cytochemically in single GM 151 cells and the mean values converted to nmol per h per 10^6 cells. GM 151 cells cultured alone had no detectable activity in these conditions.

* Millipore filters (0.22-μm pore diameter) were fitted to 20-mm diameter glass rings. The chambers, open at the top and raised on three polythene legs (2 mm high, 4 mm diameter), were placed on top of growing GM 151 monolayers, and lymphocytes or enzyme were added to the chambers. In the reverse type of co-culture, the fibroblasts were first grown on the filter inside the chamber, and the lymphocytes then placed outside in the dish. The results were identical in both culture conditions.

whether direct cell-to-cell contact was necessary, both types of cell were cultured in the same dish but were separated by Millipore filters. In these conditions, GM 151 cells did not acquire any measurable β-glucuronidase activity (Table 2). We demonstrated that free β-glucuronidase could diffuse through the Millipore membrane and be taken up by GM 151 cells on the other side by placing partially purified enzyme from 3T3 fibroblasts inside the chambers. After 24 h, enzyme uptake was 50% of that observed when the enzyme was added directly to the medium, possibly because of some adsorption of enzyme to the membrane or reduced diffusion through it. When lymphocytes and enzyme were placed in the chambers together, the uptake was again 50% of controls, showing that the lymphocytes themselves had no influence on diffusion of enzyme.

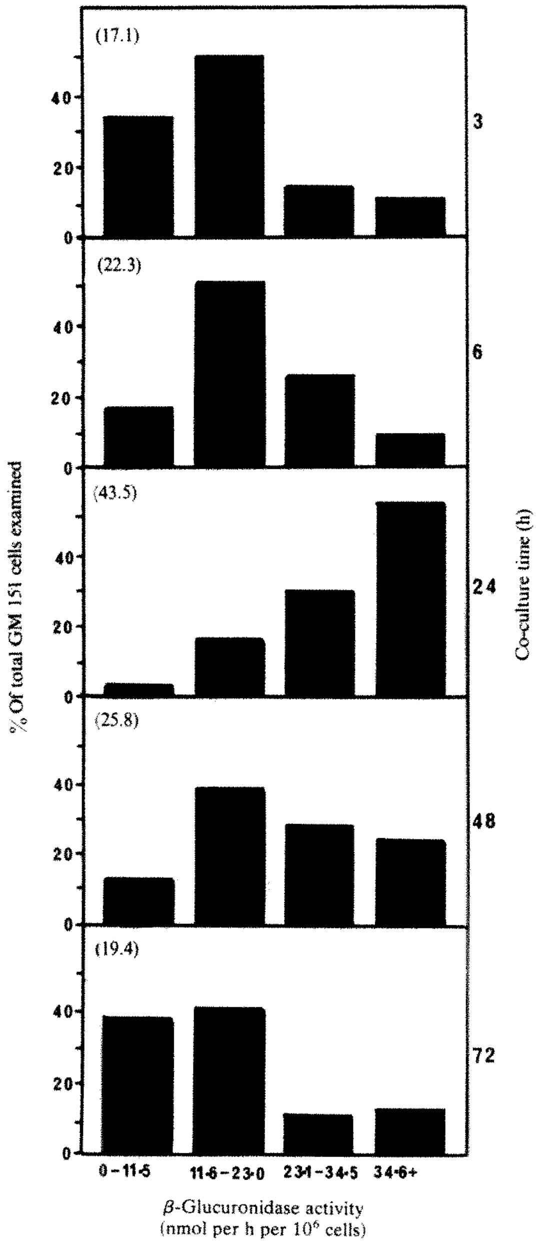


Fig. 2 Changes in β-glucuronidase activity in GM 151 cells during co-culture with lymphocytes. Monolayers of 10^5 GM 151 cells were grown directly in 35-mm culture dishes and the co-culture initiated by the addition of 10^7 lymphocytes. At the end of the co-culture periods (3–72 h), non-adherent lymphocytes were removed with PBS (37 °C), and the adherent cells trypsinized (0.25% trypsin for 10 min at 37 °C). After neutralization with serum, the cells were centrifuged, washed with PBS, and incubated with naphthol AS-BI-β-D-glucuronide substrate for 3 h at 37 °C. The cells were centrifuged, again washed with PBS, coupled with fast blue RR, washed again and resuspended in PBS. Aliquots of the mixed cell suspension were air dried on slides, then washed in methanol and water, and re-dried and mounted. GM 151 cells were readily distinguishable by their shape and larger size. The absorption densities (at 600 nm) of 30–100 individual fibroblasts were determined on a Vickers microdensitometer M-85 using a B5 mask. Results are shown as the proportion of GM 151 cells measured within a designated range of activity, which was calculated from the absorption densities using the standard curve shown in Fig. 3. Numbers in parentheses are the arithmetic means of the single-cell β-glucuronidase activities, expressed as nmol per h per 10^6 cells. Note the gaussian distribution of β-glucuronidase activity throughout the co-culture period, rather than two distinct populations of fibroblasts, one positive and the other negative for enzyme.

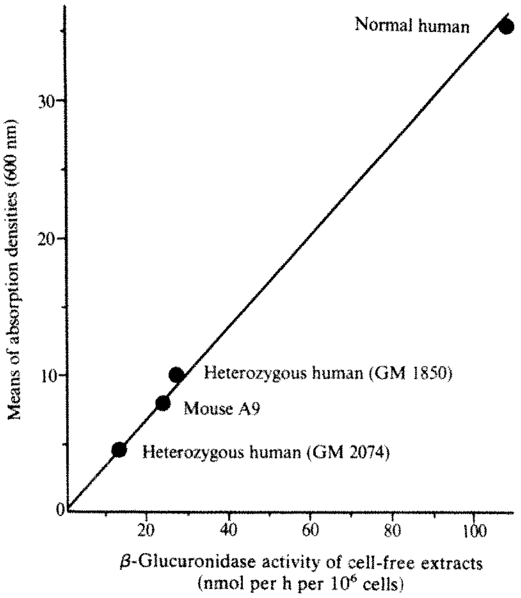


Fig. 3 Correlation of cytochemical and biochemical assays of β-glucuronidase in fibroblast cells. Mouse A9 cells, normal human fibroblasts and fibroblasts from two β-glucuronidase heterozygotes, GM 1850 and GM 2074, were grown directly on Petri dishes to late log phase. The monolayers were washed with PBS and the cells removed with trypsin. For quantitative single-cell cytochemistry, aliquots of the cell suspensions were analysed as described in Fig. 2 legend. All the cell lines had normal gaussian distributions of absorption densities, and the arithmetic means of these were calculated. For direct biochemical assay of β-glucuronidase activity, frozen and thawed extracts were prepared from aliquots of the same suspensions of fibroblasts. These were incubated for 1 h at 37 °C in 2 mM 4-methylumbelliferyl β-D-glucuronide in 0.1 M acetate buffer, *p*H 4.5 in a final volume of 0.25 ml. One ml of 0.4 M glycine buffer, *p*H 10.4, was added and the fluorescence at 448 nm measured on a Locarte fluorimeter (excitation 368 nm). β-glucuronidase activity is expressed as nm per h per 10^6 fibroblast cells used to prepare the cell-free enzyme extract. From the linear relationship between the means of the single-cell cytochemical measurements and the direct biochemical assays of β-glucuronidase in extracts of disrupted cells, a value of 2.9 nmol per h per 10^6 cells per unit density was obtained. This was used to calculate β-glucuronidase activity acquired as a result of co-culture with lymphocytes, expressed as nmol per h per 10^6 GM 151 cells.

Identical results were obtained with fibroblasts from a different patient with mucopolysaccharidosis type VII (GM 121). These findings indicate that the acquisition of β-glucuronidase activity was not due to the release of enzyme from lymphoid cells followed by its uptake by deficient fibroblasts. Thus although the precise nature of the interaction between lymphocytes and fibroblasts is unknown, direct cell-to-cell contact seems to be essential for the acquisition of enzyme activity. The present results might also explain the naturally occurring exchanges of β-glucuronidase observed in tetraparental mice[24], the presence of α-mannosidase activity in chimaeric mannosidosis calves[25], and perhaps also the transient clinical benefits following leukocyte infusion in Sanfilippo's[26] and Hunter's[27] syndromes.

We thank Miss Stephanie Diment for β-glucuronidase from 3T3 cells, Mrs R. Smith and Mrs C. Clifton for technical assistance, and Dr Chayen of the Kennedy Institute for the microdensitometer facilities. We acknowledge the supply of the

deficient fibroblasts by the Human Genetic Mutant Cell Repository.

Received 5 December 1980; accepted 17 February 1981.

1. Fratantoni, J. C., Hall, C. W. & Neufeld, E. F. *Science* **162,** 570–572 (1968).
2. Lagunoff, D., Nichol, D. M. & Pritzl, P. *Lab. Invest.* **29,** 449–453 (1973).
3. Brot, E. F., Glaser, J. H., Roozen, K. J. & Sly, W. S. *Biochem. biophys. Res. Commun.* **57,** 1–8 (1974).
4. Hultberg, B. & Sjöblad, S. *Clin. chim. Acta* **80,** 79–86 (1977).
5. Dean, M. F. *et al. Nature* **261,** 323–325 (1976).
6. Dean, M. F. *et al. J. clin. Invest.* **63,** 138–145 (1979).
7. Gibbs, D. A. *et al.* in *Birth Defects Original Article Series* Vol. 16 (ed. Desnick, R. J.) 457–474 (March of Dimes Birth Defects Foundation, Paris, 1980).
8. Yoffey, J. M. & Courtice, F. C. *Lymphatics, Lymph and the Lymphomyeloid Complex* (Academic, London, 1970).
9. Chang, T. W., Celis, E., Eisen, H. N. & Solomon, F. *Proc. natn. Acad. Sci. U.S.A.* **76,** 2917–2921 (1979).
10. Harris, G., Olsen, I. & Furner, P. *Differentiation* **12,** 1–13 (1978).
11. Hall, C. W., Cantz, M. & Neufeld, E. F. *Archs Biochem. Biophys.* **155,** 32–38 (1973).
12. Skudlarek, M. D. & Swank, R. T. *J. biol. Chem.* **254,** 9939–9942 (1979).
13. Kaplan, A., Achord, D. T. & Sly, W. S. *Proc. natn. Acad. Sci. U.S.A.* **74,** 2026–2030 (1977).
14. Sando, G. N. & Neufeld, E. F. *Cell* **12,** 618–627 (1977).
15. Ullrich, K., Mersmann, G., Weber, E. & Von Figura, K. *Biochem. J.* **170,** 643–650 (1978).
16. Julians, R. L. *et al. Expl Cell Res.* **120,** 63–72 (1979).
17. Gonzalez-Noriega, A. & Sly, W. S. *Biochem. biophys. Res. Commun.* **85,** 174–182 (1978).
18. Chern, C. J. *Proc. natn. Acad. Sci. U.S.A.* **74,** 2948–2952 (1977).
19. Bell, C. E., Sly, W. S. & Brot, E. E. *J. clin. Invest.* **59,** 97–105 (1977).
20. Hauser, P. & Vaes, G. *Biochem. J.* **180,** 249–251 (1979).
21. Cohen, S., Pick, E. & Oppenheim, J. J. (eds) *Biology of Lymphokines* (Academic, London, 1979).
22. Wahl, S. M., Wahl, L. M. & McCarthy, J. B. *J. Immun.* **121,** 942–946 (1978).
23. Waksman, B. H. *Pharmac. Ther.* **2,** 623–672 (1978).
24. Feder, N. *Nature* **263,** 67–69 (1976).
25. Jolly, R. D. *et al. Pediat. Res.* **10,** 219–224 (1976).
26. Moser, H. W. *et al. Archs Neurol.* **31,** 329–337 (1974).
27. Knudson, A. G., Di Ferrante, N. & Curtis, J. E. *Proc. natn. Acad. Sci. U.S.A.* **68,** 1738–1741 (1971).
28. Olsen, I. *et al.* (in preparation).

Diazepam induces mitotic arrest at prometaphase by inhibiting centriolar separation

Leif C. Andersson*†, Veli-Pekka Lehto†, Svante Stenman†, R. Andrew Badley‡ & Ismo Virtanen†

* Transplantation Laboratory and † Department of Pathology,
University of Helsinki, SF 00290 Helsinki 29, Finland
‡ Unilever Research Laboratories, Colworth House, Sharnbrook,
Bedford MK44 1LO, UK

We report here that we have found an accumulation of condensed mitotic figures in cultures of various human haematopoietic cell lines treated with diazepam (40–80 μg ml^{-1}). The mode of mitosis arresting activity of diazepam was analysed with cultured human fibroblasts. Unlike conventional anti-mitotic drugs, diazepam did not affect the microtubular integrity at the concentrations used. Instead, it seems to block mitosis by inhibiting the deviation of centrioles in the prometaphase.

The different appearances of demecolcine-arrested and diazepam-arrested mitoses are seen in Fig. 1, which shows cytocentrifuged cell smears from cultures of the human hairy cell leukaemia line, JOK-1[1]. The mitotic arrest seems to be complete as the same frequency of mitotic figures was seen after 6 and 18 h in parallel cultures treated with 5 μg ml^{-1} demecolcine or 50 μg ml^{-1} diazepam.

The mechanism of the diazepam-induced mitotic arrest was further studied with cultured human fibroblasts. The mitotic apparatus of untreated fibroblasts in metaphase showed a typical spindle-shaped structure in indirect tubulin-specific fluorescence (IIF) (Fig. 2*a*) and in double-fluorescence staining for DNA, a bright plate of condensed chromosomes in the middle of the spindle (Fig. 2*b*). Non-dividing fibroblasts in the same cultures showed a fibrillar tubulin-specific staining throughout the cytoplasmic domain (Fig. 3*a*). Fibroblasts cultured in the presence of demecolcine (0.2 μg ml^{-1}) for 6 h showed only a diffuse tubulin-specific fluorescence both in interphase cells (Fig. 3*b*) and in mitotic cells (Fig. 2*c*) with dispersed chromosomes (Fig. 2*d*).

A distinct accumulation of mitotic cells (up to 25–30%) with condensed chromosomes was seen in fibroblast cultures treated with 80 μg ml^{-1} diazepam for 12 h. When the mitotic cells were stained for tubulin by the IIF technique, an asterisk-like array of microtubules was seen directed towards the centre of the condensed chromosomal plate (Fig. 2*e, f*). A distinct pair of centrioles could be seen in the centre of the microtubule asterisk (Fig. 2*e*, inset). Cells in telophase stage, with a completed chromosome separation, were not observed in diazepam-treated cultures, and the appearance of microtubules in non-

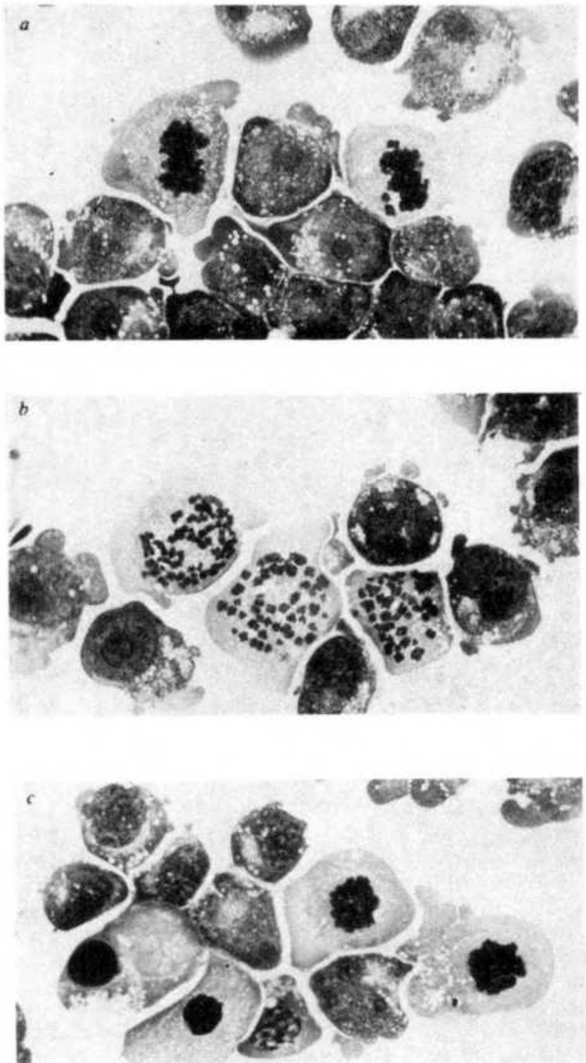


Fig. 1 May–Grünwald–Giemsa stained smears of JOK-1 cells cultivated in RPMI-1640 medium supplemented with 10% fetal calf serum (FCS). *a*, Control culture; *b*, culture treated with 0.5 μg ml^{-1} demecolcine for 6 h; *c*, culture treated with 50 μg ml^{-1} diazepam for 6 h. Diazepam (Orion) was dissolved in dimethyl sulphoxide (DMSO). The diazepam-treated cultures contained 0.04–0.08% DMSO. DMSO at these concentrations did not measurably interfere with cell growth. The frequency of mitotic figures at 6 and 18 h after initiation of cultures were 1.5 and 2.5% in control cultures, 6 and 23% in demecolcine-treated cultures, and 11.5 and 25.5% in diazepam-treated cultures. ×800.

0028–0836/81/210247—02$01.00

mitotic cells was unaltered (Fig. 3). This is different from the effect seen with conventional anti-mitotic drugs (for example, demecolcine, vinblastine sulphate), which arrest cultured cells at mitosis by disrupting the microtubules of the mitotic apparatus[2,3], as is also indicated by the diffuse staining in IIF with anti-tubulin antibodies[3] and dispersion of chromosomes.

Clark and Ryan recently reported that tranquillizers of the benzodiazipine group inhibit the proliferation of 3T3 fibroblasts in pre-S stage[4]. Our findings indicate that diazepam specifically introduces a block in the prometaphase by an unique mechanism: without affecting the microtubular integrity the centriolar separation is inhibited, giving rise to an asterisk-like mitotic apparatus. Interestingly, such a variant mitotic apparatus has also recently been described to occur at low frequency during spontaneous mitosis of PTK_2 cells[5].

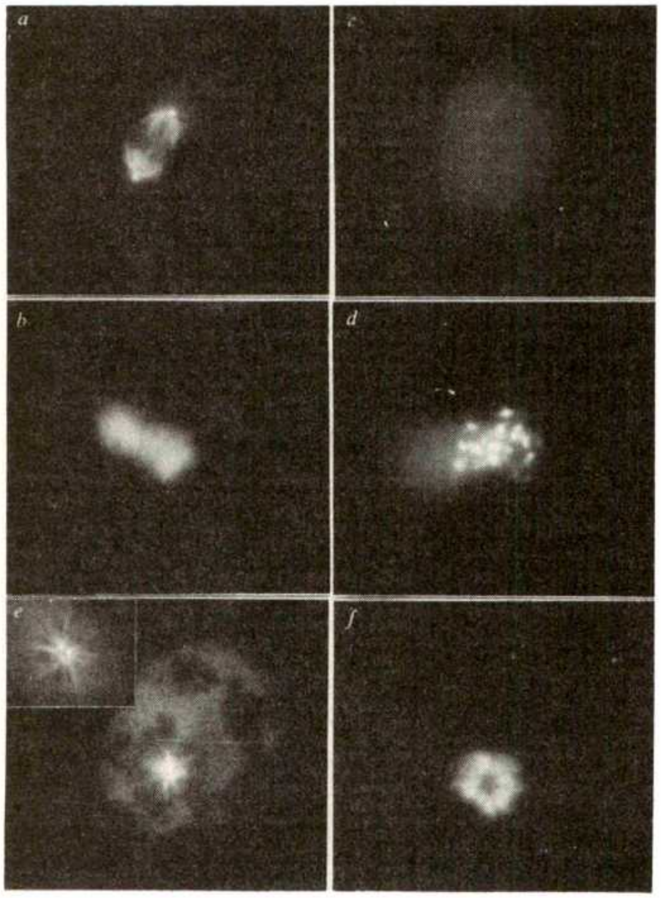


Fig. 2 Human embryonal fibroblasts were cultured in RPMI-1640 medium plus 10% FCS. For IIF microscopy the cells, cultured on small glass coverslips, were fixed in methanol for 10 min at −20 °C. Microtubules were visualized using rabbit antibodies against tubulin from calf brain[6,7]. Fluorescein isothiocyanate-coupled anti-rabbit IgG (Cappel) was used as a second antibody. Chromosomes were visualized by a double-fluorescence technique using the Hoechst no. 33258 fluorochrome[8]. In untreated fibroblasts at metaphase a bright spindle-like tubulin-specific fluorescence is seen in *a*. Note the bright spots, representing centrioles, at the poles of the spindle. In double-fluorescence with the Hoechst fluorochrome (*b*) the chromosomes were seen as a bright plate in the middle of the spindle. When fibroblasts were cultured in the presence of 0.2 $\mu g\ ml^{-1}$ demecolcine for 6 h, rounded mitotic cells accumulated in the cultures. In such cells only a diffuse tubulin-specific staining (*c*) and scattered chromosomal figures (*d*) were seen. When human fibroblasts were cultured in the presence of diazepam (80 $\mu g\ ml^{-1}$ for 12 h), IIF microscopy with anti-tubulin antibodies revealed an asterisk-like mitotic apparatus (*e*) directed towards the centre of the condensed chromosomal plate as seen in double fluorescence microscopy with the DNA-fluorochrome (*f*). The centrioles, visualized as bright spots with tubulin-antibodies, were seen as a distinct doublet in the centre of the microtubule asterisk (*e*, inset). ×680.

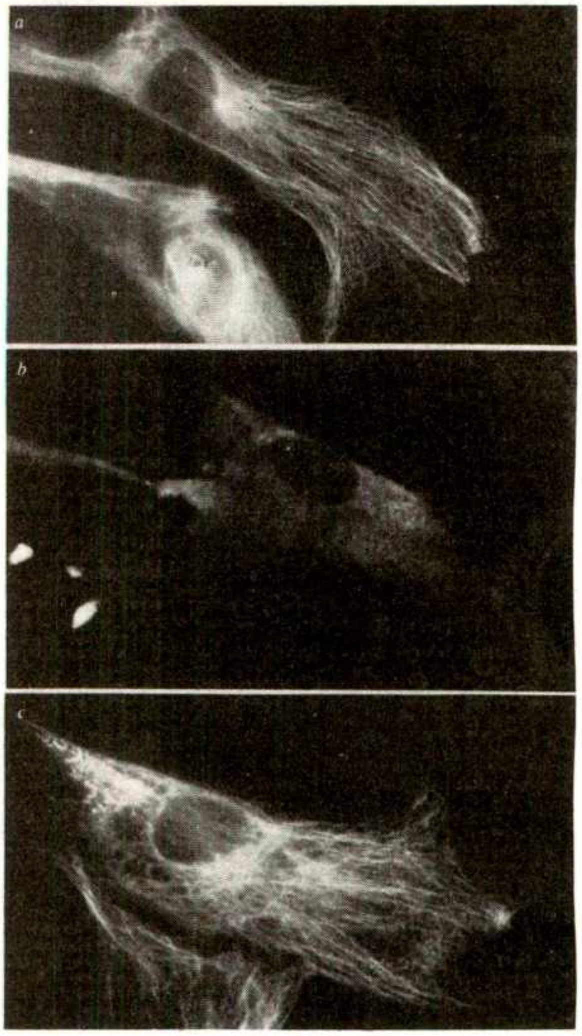


Fig. 3 In normal non-dividing cultured fibroblasts microtubules were visualized by the IIF technique as fibrillar cytoplasmic arrays throughout the whole cytoplasmic domain (*a*). After addition of demecolcine only a diffuse cytoplasmic tubulin-specific fluorescence was seen in the treated cells (*b*). In fibroblasts cultured in the presence of diazepam (80 $\mu g\ ml^{-1}$), normal fibrillar arrays of microtubules were seen (*c*). ×400.

We thank Ms P. Nieminen and Ms R. Taavela for technical assistance and Ms L. Saraste for secretarial help. This study was supported by grants from the Finnish MRC, the Finnish Cancer Society, the Finska Läkaresällskapet, Helsinki, Finland and by grant no. 1 RO1 CA26294-01A1 PBC from the NIH.

Received 26 January; accepted 16 March 1981.

1. Andersson, L. C. & Gahmberg, C. G. *Proc. 3rd Bristol-Myers Symp. Cancer Res.*, 1981 (in the press).
2. McIntosh, J. R. in *Microtubules* (eds Roberts, K. & Hyams, J. S.) 315–358 (Academic, London, 1979).
3. Cande, W. Z., Lazarides, E. & McIntosh, J. R. *J. Cell Biol.* **72,** 552–567 (1977).
4. Clarke, G. D. & Ryan, P. J. *Nature* **287,** 160–161 (1980).
5. Aubin, J., Osborn, M. & Weber, K. *J. Cell Sci.* **43,** 177–194 (1980).
6. Badley, R. A. *et al. Expl Cell Res.* **117,** 231–244 (1978).
7. Virtanen, I., Lehto, V.-P., Lehtonen, E. & Badley, R. A. *Eur. J. cell. Biol.* **23,** 80–84 (1980).
8. Russell, W. C., Newman, C. & Williamson, D. H. *Nature* **253,** 461–462 (1975).

Intermediate filaments in 3T3 cells collapse after intracellular injection of a monoclonal anti-intermediate filament antibody

Michael W. Klymkowsky

MRC Neuroimmunology Project, Department of Zoology and
Comparative Anatomy, University College London,
London WC1E 6BT, UK

The function of intermediate filaments within cells is unclear, largely because there are no known reagents or physical conditions which specifically affect their organization (see ref. 1 for review). I report here that the intracellular injection of a monoclonal IgG1 antibody specific for intermediate filaments[2] causes the rapid and reversible collapse of all intermediate filaments within 3T3 mouse fibroblasts while leaving the organization of microfilament bundles and microtubules apparently unaltered. Direct observation of cells undergoing intermediate filament collapse indicates that such collapse occurs without obvious effects on cell morphology or motility.

Monoclonal antibody against an intermediate filament antigen (IFA)[2] was purified from ascites fluid raised in *nu/nu* mice on a column of *Staphylococcus* A–Sepharose (Pharmacia), concentrated by vacuum dialysis, dialysed into injection buffer (114 mM KCl, 20 mM NaCl, 3 mM $MgCl_2$, 3 mM NaH_2PO_4, *p*H 7 with KOH (ref. 3)) and spun at 100,000*g* for 30 min before use. About 2×10^{-14}l (ref. 4) of antibody at a concentration of 3 mg ml^{-1} was injected into 3T3 (clone A31) cells growing on glass coverslips for 20–48 h. During and after injection, cells were maintained at 37°C until they were fixed after various periods of time with methanol for 10 min at −20°C. The cells were then stained sequentially with fluorescein-coupled goat anti-mouse immunoglobulin (G-anti-MIg-F) to visualize microinjected antibodies and by rabbit anti-tubulin[5], rabbit anti-cytoplasmic myosin[6] or more monoclonal anti-IFA antibodies followed by rhodamine-coupled goat anti-rabbit immunoglobulin (G-anti-RIg-Rd) or G-anti-MIg-Rd. Appropriate control experiments indicated that all reagents used were specific (data not shown). After washing, the coverslips were mounted in glycerol and viewed with a Zeiss Universal microscope equipped with phase contrast and epifluorescence optics.

When fixed within 1 h of injection and stained with G-anti-MIg-F, many injected cells were seen to have their intermediate filaments collapsed within a bundle near or around the cell's nucleus. When such cells were exposed to more anti-IFA antibody followed by G-anti-MIg-Rd, no intermediate filament staining was observed (Fig. 1) outside the collapsed bundle. On exposure of these to either anti-tubulin or anti-cytoplasmic myosin antibodies, followed by G-anti-RIg-Rd, most microtubules and microfilament bundles were unaffected by the collapse of intermediate filaments (Fig. 2). Moreover, the collapsed intermediate filaments did not label with either the anti-tubulin or anti-myosin antibodies (Fig. 2). Cells injected with tetramethylrhodamine-coupled bovine serum albumin or antibody directed against cytoplasmic myosin[6] never showed collapsed intermediate filaments, excluding the possibility that filament collapse was a response to injection.

The collapse of intermediate filaments by injected anti-IFA antibody was reversible, as demonstrated by time-course experiments such as those outlined in Table 1. Filament collapse was most marked at 1–3 h post-injection when more than 70% of injected cells had collapsed intermediate filaments; with time the percentage decreased and the extent of filament collapse in individual cells diminished until, at 16 h after injection, many cells had fully extended intermediate filaments, although these were still decorated with anti-IFA antibody. Autophagy of injected antibody[7] might have reduced the intracellular concentration of antibody to a level insufficient to maintain filament collapse.

Table 1 Time course of collapse of intermediate filaments by injected anti-IFA antibody

Time after injection (h)	Cells recovered	Cells with collapsed intermediate filaments
1	90	79 (88)
3	89	63(71)
6	94	45(48)
16	68	16 (24)

100 cells were injected on each coverslip and after staining, the entire coverslip was scanned. The fact that less than 100% of the cells were 'recovered' reflects both cell death due to injection trauma and the inability to recognize cells due either to reduction of staining level to near background, or interfering flaws in mounting. Numbers in parentheses are the percentage of recovered cells found to be collapsed.

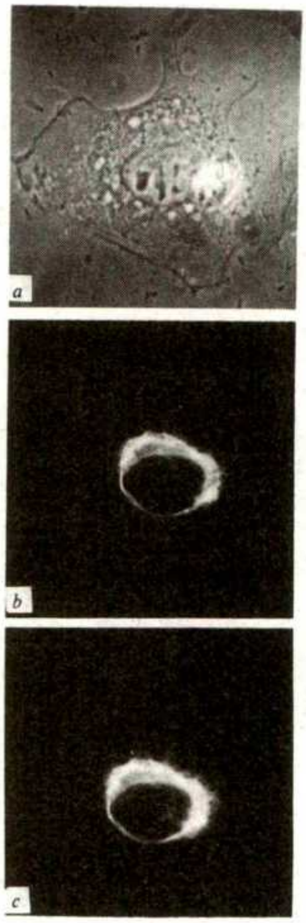


Fig. 1 Phase (*a*) and immunofluorescence of 3T3 cell injected with anti-IFA 1 h previously, fixed and then stained sequentially with G-anti-MIg-F, anti-IFA and G-anti-MIg-Rd. Intermediate filaments are collapsed around the nucleus (*b*, fluorescein optics) and secondary staining with anti-IFA shows no intermediate filaments outside the collapsed bundle (*c*, rhodamine optics). Cells were injected in HEPES-buffered minimal Eagle's media (MEM) + 10% fetal calf sera (FCS) + 50 μg ml^{-1} gentamycin and then returned to bicarbonate-buffered MEM + 10% FCS and kept in 8% CO_2 at 37 °C until fixation. ×700.

0028–0836/81/210249—03$01.00

To study the effects of intermediate filament collapse on cell morphology and movement, purified tetramethylrhodamine-conjugated anti-IFA (anti-IFA-Rd) was injected into 3T3 cells and the cells were followed by time-lapse video using a Zeiss inverted microscope equipped with phase contrast and epifluorescence optics coupled to a silicon-intensified target video camera (RT Laboratories) and video recording system. Cells were examined using epifluorescence optics at the end of the observation period. Injected anti-IFA antibody spread rapidly throughout the cell and decorated intermediate filaments. Most cells retained their pre-injection morphologies, although a small proportion underwent a rapid but reversible contraction; as this type of acute shape change was just as frequently observed after injection of buffer alone, it was probably due to the injection process itself. Most injected cells showed intermediate filament collapse within 1–3 h and the morphologies and movements of the majority of such cells could not be distinguished from uninjected cells in the same field. An example of intermediate filament collapse in two injected cells is shown in Fig. 3. Because all the cells were continuously moving and changing shape, I cannot exclude the possibility that the collapse of intermediate filaments was associated with subtle changes in cell shape or movement.

The collapse of intermediate filaments induced by anti-IFA antibody was not restricted to 3T3 cells. The intermediate filaments of baby hamster kidney-21 (BHK-21) cells and secondary chick heart fibroblasts also underwent intermediate filament collapse after antibody injection and retained substantially unaltered morphologies. Very large cells often failed to show antibody-induced filament collapse. Whether this is the result of a failure to inject sufficient quantities of the antibody or is due to a true difference in the intermediate filaments of such cells is unclear.

The mechanism of anti-IFA antibody-induced collapse of intermediate filaments is unknown. The antibody has been shown to bind to vimentin, the major protein in intermediate filaments in 3T3 cells[2]. However, the simple binding of an antibody to filaments within living cells does not invariably change the organization of those filaments. For example, rabbit anti-cytoplasmic myosin antibodies injected into 3T3 cells decorated microfilament bundles but had no appreciable effect on the distribution of either microfilament bundles or intermediate filaments within the cells (data not shown). Moreover, various monoclonal antibodies which react with tonofilaments in PtK_2 cells[8] decorate these filaments after injection but do not change their distribution; anti-IFA antibodies, on the other hand, collapse tonofilaments in PtK_2 cells after injection[9].

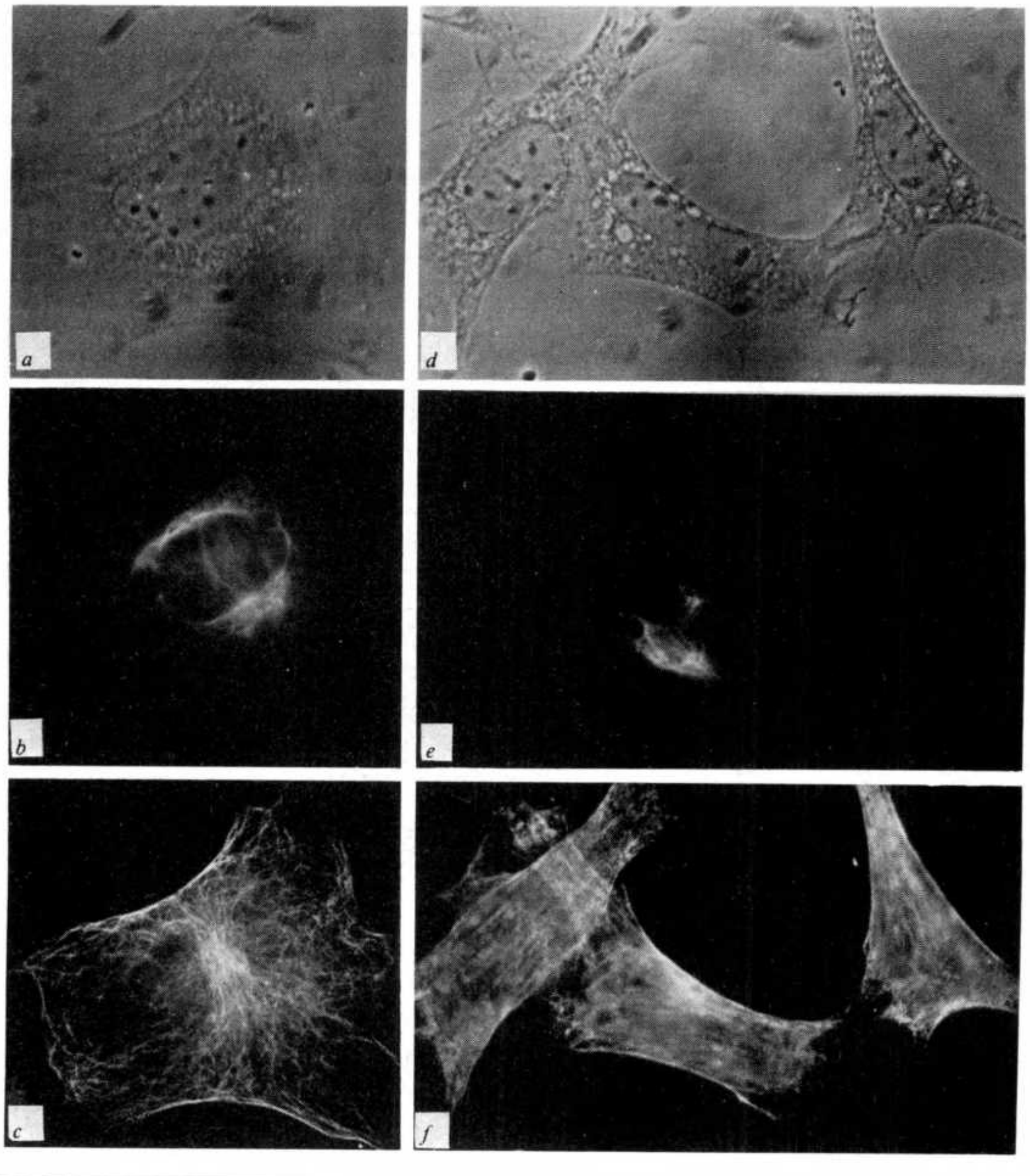


Fig. 2 Phase (*a*, *d*) and immunofluorescence of anti-IFA injected 3T3 cells (1–1·5 h before fixation) that have been stained sequentially with G-anti-MIg-F, rabbit anti-tubulin (*a*, *b*, *c*) or rabbit anti-myosin (*d*, *e*, *f*) and G-anti-RIg-Rd. Both rabbit anti-tubulin and myosin and anti-rabbit-Ig-Rd were diluted into purified mouse immunoglobulin (final concentration ~2 mg ml^{-1}) to block possible cross-reactivity with G-anti-MIg-F. G-anti-RIg-Rd had been pre-adsorbed with mouse immunoglobulin. Intermediate filaments (*b*, *e*, fluorescein optics) are collapsed by injected anti-IFA but microtubule and microfilament bundle organization appear unaltered (*c*, *f*, rhodamine optics). ×1,060.

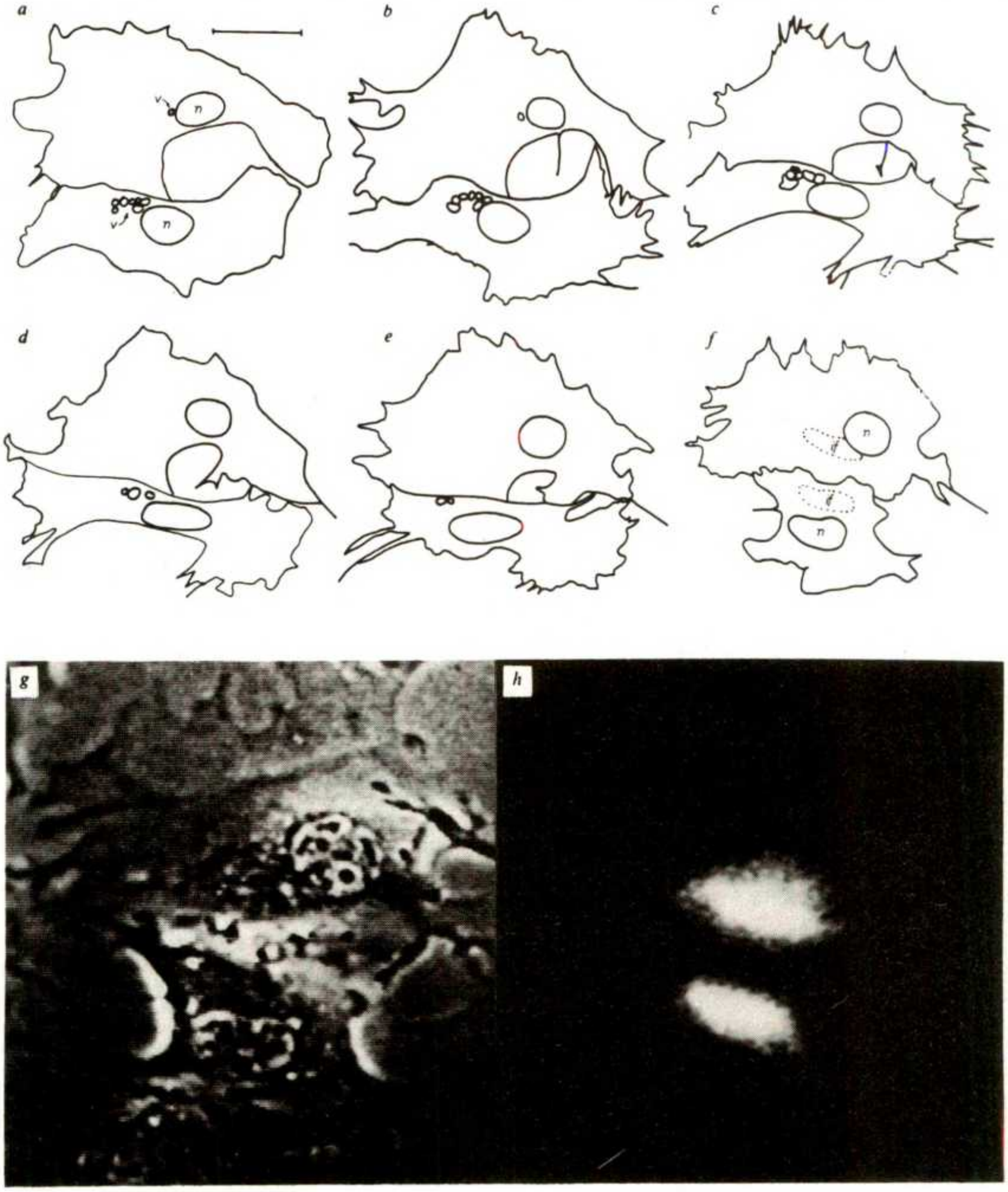


Fig. 3 Two cells injected with anti-IFA-Rd were followed continuously by time-lapse video. Outlines of the cells were traced from the video images before (*a*) and 2.5 min (*b*), 17 min (*c*), 36.5 min (*d*), 59.5 min (*e*), and 76 min (*f*) after injection. The cells were viewed under rhodamine optics 76 min after injection, revealing their collapsed intermediate filaments (region defined by dotted boundary and 'if' in *f*). The phase contrast (*g*) and rhodamine optics images (*h*) of the cells 76 min after injection as photographed from the video monitor are also shown. The cells' nuclei 'n' are labelled (*a*, *f*) and vacuoles 'v' (in *a*), which subsequently disappear. Similar shape changes to those seen in these two injected cells are frequently seen in normal, uninjected 3T3 cells over this period of time. Cells were maintained at 37 °C in HEPES-buffered MEM + 10% FCS + 50 μg ml^{-1} gentamycin throughout. To minimize evaporation of media the culture dish was overlaid with paraffin oil. Cells survived and divided apparently normally in these conditions for longer than 24 h. Scale bar, 50 μm.

Increasing indirect evidence suggests that intermediate filaments and microtubules are functionally and perhaps physically associated with each other[10]. The effects of microtubule depolymerization, whether by specific drugs or physical conditions, results in the eventual collapse of intermediate filaments, not unlike that seen after injection of anti-IFA, although with a much slower time course[1]. The results presented here indicate that whereas the extended organization of intermediate filaments may well depend on microtubules, the converse is clearly not the case. Apparently, intermediate filaments can be specifically withdrawn from the cytoskeletal network of the cell without obvious effects on the organization of microfilament bundles or most microtubules, or on gross cell morphology or movement. These findings suggest that intermediate filaments may have a more subtle role than previously believed.

While this manuscript was in preparation, similar results were reported in abstract form by Lin and Feramisco[11] using a monoclonal IgM antibody directed against a 95,000 molecular weight protein associated with intermediate filaments.

I thank members of the MRC Neuroimmunology Project, especially Dr R. Pruss for anti-IFA antibody and Drs M. Raff, M. Noble, E. B. Lane and S. Bevan, for their stimulating suggestions and encouragement, Dr T. Preston for the use of his micromanipulator, and Drs P. Sheterline and J. Fallon for gifts of anti-tubulin antisera and anti-cytoplasmic myosin antibodies respectively. I am a Muscular Dystrophy Association of America postdoctoral fellow.

Received 8 December 1980; accepted 24 February 1981.

1. Lazarides, E. *Nature* **283,** 249–256 (1980).
2. Pruss, R. M., Mirsky, R., Raff, M. C., Thorpe, R. & Anderton, B. H. *J. Cell Biol.* **87,** 187a (1980).
3. Stacey, D. W. & Allfrey, V. G. *Cell* **9,** 725–732 (1976).
4. Graessman, M. & Graessman, A. *Proc. natn. Acad. Sci. U.S.A.* **73,** 366–370 (1976).
5. Sheterline, P. *Expl Cell Res.* **115,** 460–464 (1978).
6. Fallon, J. R. & Nachmias, V. *J. Cell Biol.* **87,** 237–247 (1980).
7. Stacey, D. W. & Allfrey, V. G. *J. Cell Biol.* **75,** 807–817 (1977).
8. Lane, E. B. *J. Cell Biol.* (submitted).
9. Klymkowsky, M. W. & Lane, E. B. (in preparation).
10. Geiger, B. & Singer, S. D. *Proc. natn. Acad. Sci. U.S.A.* **77,** 4769–4773 (1980).
11. Lin, J. C. C. & Feramisco, J. R. *J. Cell Biol.* **87,** 181a (1980).

Regulation of human platelet myosin light chain kinase by the catalytic subunit of cyclic AMP-dependent protein kinase

D. R. Hathaway*, C. R. Eaton & R. S. Adelstein†

Section on Molecular Cardiology, Cardiology Branch, National
Heart, Lung, and Blood Institute, Building 10, Room 7B–15,
National Institutes of Health, Bethesda, Maryland 20205, USA

A rise in the concentration of cyclic AMP inhibits secretion in human platelets[1–3]. In mammalian cells the mechanism of action of cyclic AMP involves activation of a protein kinase and results in protein phosphorylation[4]. It has previously been shown that myosin light chain kinase purified from smooth muscle is a substrate for the catalytic subunit of cyclic AMP-dependent protein kinase[5] and that phosphorylation of smooth muscle myosin kinase results in a decrease in its activity[6]. Here we present evidence that platelet myosin kinase is a substrate for the catalytic subunit of protein kinase and that phosphorylation of this myosin kinase, isolated from a non-muscle cell, decreases the activity of this enzyme. Dephosphorylation of platelet myosin kinase by a purified phosphatase[7] restores its original activity. A decrease in myosin kinase activity in platelets would increase the relative amount of unphosphorylated myosin, which unlike phosphorylated myosin, cannot interact with actin[8,9]. These findings suggest a mechanism by which cyclic AMP might modulate actin–myosin interaction in platelets and other non-muscle cells.

The platelet myosin kinase used was purified by affinity chromatography using a column of calmodulin–Sepharose 4B (refs 10, 11). Figure 1 shows the time course of incubating purified platelet myosin kinase, $[\gamma\text{-}^{32}P]$ATP and the catalytic subunit of protein kinase in the presence and absence of protein kinase inhibitor. In the presence of the catalytic subunit, ^{32}P was rapidly incorporated into platelet myosin kinase, whereas incubation in the presence of the protein kinase and its specific inhibitor decreased the amount of ^{32}P incorporation by ~90%. Incubation in the absence of protein kinase showed only minimal amounts of incorporation of ^{32}P after 1 h. A time course experiment carried out in the absence of protein kinase showed no evidence for autophosphorylation during earlier times (not shown).

To verify that platelet myosin kinase served as the phosphate acceptor, a sample of myosin kinase that had been incubated for 45 min with $[\gamma\text{-}^{32}P]$ATP and protein kinase was subjected to SDS-polyacrylamide gel electrophoresis (SDS-PAGE). The major protein peak ($M_r = 100{,}000$), corresponding to myosin kinase, coincided with the only peak of protein-bound ^{32}P. When PAGE was carried out in the absence of SDS, a single peak of myosin kinase activity, which contained all the ^{32}P, was eluted from the gel (data not shown).

The stoichiometry of phosphorylation was determined both by Millipore filtration of phosphorylated myosin kinase following precipitation with trichloroacetic acid and by quantifying the amount of protein and ^{32}P migrating with myosin kinase in SDS-PAGE. In four separate preparations an average of 1.2 mol of phosphate per mol of myosin kinase was incorporated. However, these results do not exclude the possibility that more than one site can be phosphorylated.

Figure 2 shows the effect of phosphorylation on platelet myosin kinase activity and is a double reciprocal plot of data derived from activation of myosin kinase over a range of calmodulin concentrations. Two effects of phosphorylation are observed. First, phosphorylation results in a threefold decrease in maximal enzymatic activity (from 0.148 to 0.054 nmol min^{-1}). Second, the concentration of calmodulin required for half-maximal enzyme activity is threefold greater for phosphorylated myosin kinase than for the unphosphorylated species (2.8×10^{-8} M compared with 0.9×10^{-8} M).

The effect of phosphorylating myosin kinase is reversible, as shown in Fig. 3, which compares the specific activity of unphosphorylated, phosphorylated and dephosphorylated platelet myosin kinase at two different concentrations of calmodulin. Figure 3 also shows an autoradiogram indicating the state of phosphorylation of the kinase before and after incubation with phosphatase. The enzyme used to dephosphorylate myosin kinase was the purified phosphatase I prepared from turkey gizzards[7], which was free of protease activity. The ability of smooth muscle phosphatase I to dephosphorylate platelet myosin kinase and restore the higher activity of the unphosphorylated enzyme is consistent with a role for this phosphorylation in regulating the activity of myosin kinase.

Phosphorylation of the 20,000-M_r light chain of platelet myosin and other non-muscle myosins has been shown to have two important effects: (1) it is required for the actin-activation of the Mg-ATPase activity of myosin[8,9,12–14], and (2) it promotes the assembly of myosin filaments in physiological conditions[15]. The enzyme catalysing phosphorylation of platelet myosin has been purified and shown to have an absolute requirement for

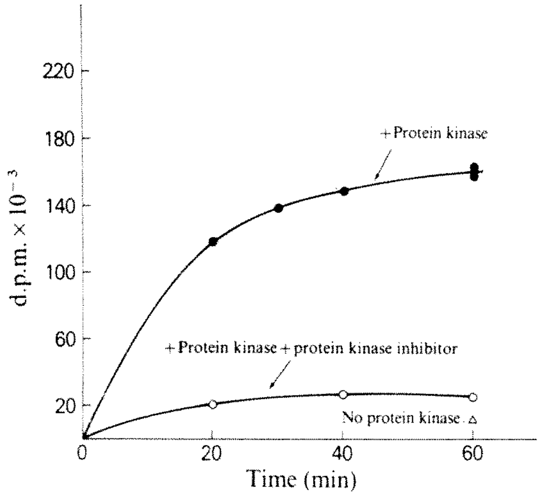


Fig. 1 Time course of phosphorylation of platelet myosin light chain kinase by the catalytic subunit of cyclic AMP-dependent protein kinase in the presence and absence of the specific inhibitor of protein kinase. Myosin light chain kinase was purified as described by D.R.H. and R.S.A.[10] except that the crude extract was applied to a 20-ml column of AH-Sepharose 4B and eluted with 0.5 M NaCl, replacing the DEAE-Sephacel chromatography, and Sephacryl S-300 was used rather than S-200. The purified kinase (5×10^{-8} M) was incubated at 23 °C in 50 mM Tris-HCl (*p*H 7.5), 5 mM $MgCl_2$, 25 μM $[\gamma\text{-}^{32}P]$ATP (2×10^7 c.p.m. nmol^{-1}) and 2×10^{-7} M cyclic AMP-dependent protein kinase (a gift of Dr E. H. Fischer). This was done in the presence (○) and absence (●) of the specific inhibitor of cyclic AMP-dependent protein kinase[18] (a gift from Dr L. Rappaport). Time points were taken at 20, 30, 40 and 60 min, and the amount of ^{32}P incorporated into the myosin kinase was determined by Millipore filtration and liquid scintillation counting of the filter disks[12]. A single time point was taken in the absence of protein kinase (△), showing no phosphorylation of the platelet myosin kinase. When histone IIA was phosphorylated at comparable concentrations using the same amount of protein kinase, the time course of phosphorylation was similar to that obtained with platelet myosin kinase. Phosphorylation of platelet myosin kinase was not dependent on Ca^{2+}.

---

* Present address: Department of Medicine, Indiana University School of Medicine, Indianapolis, Indiana 46223, USA.

† To whom correspondence should be addressed.

0028-0836/81/210252—03$01.00

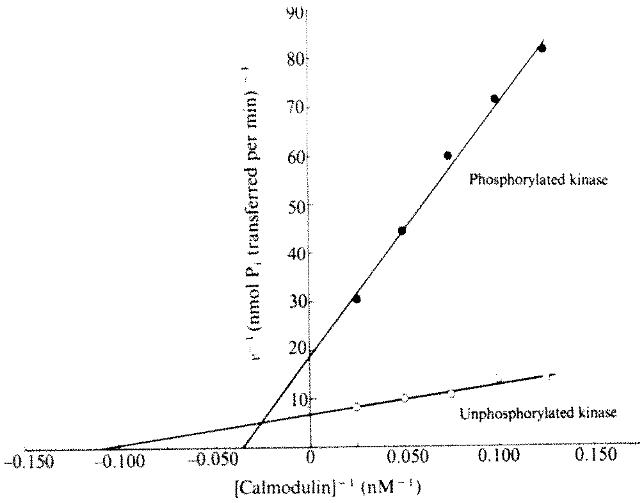


Fig. 2 Double reciprocal plot showing the effect of phosphorylation on the activity of platelet myosin light chain kinase. The myosin kinase (3×10^{-7} M) was incubated for 45 min at 23 °C with 40 mM Tris-HCl (*p*H 7.5), 5 mM $MgCl_2$ and 4×10^{-7} M cyclic AMP-dependent protein kinase, in the presence (●) and absence (○) of 50 μM ATP. The resulting phosphorylated and unphosphorylated samples of myosin kinase were then each assayed over a range of calmodulin concentrations from 9×10^{-9} M to 4×10^{-8} M. Calmodulin was prepared from bovine brain by the method of Klee[19]. The concentration of platelet myosin kinase was 3.3×10^{-8} M. The assays were carried out for 1 min at 23 °C in 50 mM Tris-HCl (*p*H 7.3), 10 mM $MgCl_2$, 0.1 mM $CaCl_2$, 100 μM [γ-^{32}P]ATP (2×10^6 c.p.m. nmol^{-1}) and 25 μM smooth muscle myosin 20,000-M_r light chain. This myosin light chain fraction was prepared as outlined previously[5]. Total assay volume was 100 μl.

Ca^{2+} and the ubiquitous calcium-binding protein, calmodulin[10,11].

Previous work has shown that the enzyme catalysing myosin phosphorylation in smooth muscle can be phosphorylated by the catalytic subunit of cyclic AMP-dependent protein kinase[5]. Phosphorylation has a marked effect on the ability of the smooth muscle enzyme to bind calmodulin, but only a small effect on V_{max} (ref. 6). The results reported here for platelet myosin kinase differ in that phosphorylation of the platelet myosin kinase decreases both the ability of the enzyme to bind calmodulin and the activity of the enzyme when extrapolated to an infinite concentration of calmodulin (V_{max}). This difference in the effect of phosphorylation on the kinetic parameters of the two enzymes may simply reflect the species difference between the turkey gizzard smooth muscle kinase and that isolated from human platelets. On the other hand, it is possible that the purified platelet myosin kinase, which has a molecular weight 30,000 less than the smooth muscle kinase (100,000 compared with 130,000), has undergone partial proteolysis. Such proteolysis, which does not alter the requirement of the enzyme for calmodulin or the ability to respond to phosphorylation, could be responsible for a subtle alteration in the kinetic response. Further work will be necessary to rule out this latter possibility.

Cyclic nucleotide-mediated phosphorylation of myosin kinases may provide an important mechanism for modulating contractile activity in non-muscle cells. A rise in the intracellular concentration of cyclic AMP would result in the phosphorylation of myosin kinase and hence decrease myosin kinase activity. The net result of this decreased activity would be an inhibition of contractile activity because it would result in a relative increase in the unphosphorylated form of myosin. Indirect evidence for cyclic AMP-mediated regulation in smooth muscle cells has been provided by Kerrick *et al.*[16], who showed that addition of the catalytic subunit of cyclic AMP-dependent protein kinase to chemically skinned smooth muscle fibres resulted in a decrease in both tension and the phosphorylation of smooth muscle myosin light chains.

Daniel *et al.*[17] have shown that platelet secretion can be correlated with the phosphorylation of platelet myosin, indicating that secretion results from contraction of platelet actomyosin. This finding, together with data from physiological experiments suggesting that agents which increase cyclic AMP in platelets act to inhibit secretion[1-3], is consistent with our *in vitro* findings that phosphorylation of platelet myosin kinase by cyclic AMP-dependent protein kinase results in inhibition of myosin kinase activity. Of course, it is possible that cyclic AMP works through more than one pathway to cause a decrease in contractile activity in platelets. To confirm the physiological significance of the mechanism outlined here, future experiments must: (1) demonstrate phosphorylation of platelet myosin kinase *in vivo* in response to a rise in intracellular levels of cyclic AMP and (2) correlate this phosphorylation with inhibition of a contractile event. However, our present data suggest that phosphorylation of myosin kinase by cyclic AMP-dependent protein kinase is one mechanism by which cyclic AMP regulates contractile activity in platelets and other non-muscle cells.

We thank Mr William Anderson Jr and Mr James M. Miles for technical assistance, and Mrs Exa Murray for preparing this manuscript.

Specific activity of myosin kinase (μmol per mg per min)

	1 x 10-9M Calmodulin	3 x 10-9M Calmodulin
UNPHOSPHORYLATED	0.21	0.98
PHOSPHORYLATED	0.09	0.45
DEPHOSPHORYLATED	0.23	0.96

Fig. 3 The effect of dephosphorylation of myosin kinase activity. Platelet myosin kinase was incubated in the presence and absence of 1.25 μM catalytic subunit of cyclic AMP-dependent protein kinase. Also present in both tubes were 50 mM Tris-HCl (*p*H 7.3), 10 mM $MgCl_2$, 0.1 mM $CaCl_2$ and 20 μM [γ-^{32}P]ATP. After 30 min at 23 °C, 20 mN EDTA was added to each sample to inhibit further phosphorylation. The sample with the protein kinase was then divided into two parts, to one half of which was added 6×10^{-9} M purified smooth muscle phosphatase I[7] (a gift of Dr Mary Pato). The resulting three tubes (U = no protein kinase, no phosphatase; P = +protein kinase, no phosphatase; D = +protein kinase, +phosphatase) were incubated for 15 min at 23 °C before assay. The left-hand side is an autoradiogram showing the state of phosphorylation of the phosphorylated (P) and dephosphorylated (D) samples. There is no evidence for autophosphorylation of the myosin kinase in the unphsophorylated (U) sample. The dye front is indicated. In wells P and D, halfway between the ^{32}P band due to the myosin kinase and the dye, is a light band which is probably due to autophosphorylation of the catalytic subunit of protein kinase. The samples were assayed at the two calmodulin concentrations shown, using 1×10^{-9} M platelet myosin kinase. The assays were terminated after 15 min (1×10^{-9} M calmodulin) or 8 min (3×10^{-9} M calmodulin) at 23 °C, with other conditions as in Fig. 2 legend.

Note added in proof: Recent experiments indicate that phosphate is incorporated into two different sites in platelet myosin kinase (C.R.E. and R.S.A., unpublished), as recently reported for smooth muscle myosin kinase[6].

Received 30 October 1980; accepted 2 March 1981.

1. Haslam, R. J. & Lynham, J. A. *Biochem. biophys. Res. Commun.* **77,** 714–722 (1977).
2. Haslam, R. J. & Lynham, J. A. *Thromb. Res.* **12,** 619–628 (1978).
3. Haslam, R. J., Lynham, J. A. & Fox, J. E. B. *Biochem. J.* **178,** 397–407 (1979).
4. Krebs, E. G. & Beavo, J. A. *A. Rev. Biochem.* **48,** 923–959 (1979).
5. Adelstein, R. S., Conti, M. A., Hathaway, D. R. & Klee, C. B. *J. biol. Chem.* **253,** 8347–8350 (1978).
6. Conti, M. A. & Adelstein, R. S. *J. biol. Chem.* **256,** 3178–3182 (1981).
7. Pato, M. D. & Adelstein, R. S. *J. biol. Chem.* **255,** 6535–6538 (1980).
8. Adelstein, R. S. & Conti, M. A. *Nature* **256,** 597–598 (1975).
9. Lebowitz, E. A. & Cooke, R. *J. biol. Chem.* **253,** 5443–5447 (1978).
10. Hathaway, D. R. & Adelstein, R. S. *Proc. natn. Acad. Sci. U.S.A.* **76,** 1653–1657 (1979).
11. Dabrowska, R. & Hartshorne, D. J. *Biochem. biophys. Res. Commun.* **85,** 1352–1359 (1978).
12. Scordilis, S. P. & Adelstein, R. S. *J. biol. Chem.* **253,** 9041–9048 (1978).
13. Trotter, J. A. & Adelstein, R. S. *J. biol. Chem.* **254,** 8781–8785 (1979).
14. Yerna, M.-J., Aksoy, M. O., Hartshorne, D. J. & Goldman, R. D. *J. Cell Sci.* **31,** 411–429 (1978).
15. Scholey, J. M., Taylor, K. A. & Kendrick-Jones, J. *Nature* **287,** 233–235 (1980).
16. Kerrick, W. G. L., Hoar, P. E., Cassidy, P. S. & Malencik, D. A. *Regulation of Muscle Contraction: Excitation–Contraction Coupling* (eds Grinnell, A. D. & Brazier, M. A. B.) 227–239 (Academic, New York, 1980).
17. Daniel, J. L., Molish, I. R., Holmsen, H. & Salganicoff, L. *Cold Spring Harb. Conf. Cell Proliferation* Vol. 8 (in the press).
18. Walsh, D. A. *et al. J. biol. Chem.* **246,** 1977–1985 (1971).
19. Klee, C. B. *Biochemistry* **16,** 1017–1024 (1977).

Localized gating of egg Na$^+$ channels by sperm

Meredith Gould-Somero

Department of Biology, B-022, University of California, San Diego, La Jolla, California 92093, USA

Eggs of many species respond to sperm with a positive shift in membrane potential, the fertilization potential[1], which in several[2–5] but not all[6] species mediates a block to polyspermy. How the sperm elicits the fertilization potential is not known, although it is the earliest known response of eggs to sperm, occurring within 1–2 s of mixing the gametes. In the marine worm, *Urechis caupo*, the fertilization potential results from an increase in Na$^+$ permeability which is amplified during the first 15 s by a voltage-gated Ca^{2+} action potential[7]. The Na$^+$ channels are gated not by voltage, but by sperm, a sperm opening only a fraction of available channels[7]. The experiments reported here indicate that the open Na$^+$ channels are localized in a 'patch' near the fertilizing sperm.

Because the amplitude of the fertilization potential depends on the external Na$^+$ concentration ([Na$^+$])[7], it is possible to distinguish between a localized and a uniform distribution of open Na$^+$ channels if the concentration of Na$^+$ at the site of fertilization is altered independently of that over the rest of the egg. The experimental apparatus, described in Fig. 1 legend, consisted of a perforated nickel screen in which the egg was seated separating two chambers containing seawater of the same or different [Na$^+$]; [Na$^+$] was either 470 mM (natural seawater) or 1/10th that amount. By injecting a dense sperm suspension into this chamber, then removing the egg and measuring the area to which sperm were bound, at ×400 with an ocular micrometer, we calculated that 5–10% of the egg surface was exposed to the solution in the lower chamber. Because the resistance through the egg interior is small relative to that across the membrane, the egg is isopotential, and the electrode in the upper surface of the egg accurately records the fertilization potential initiated by a sperm at the opposite side of the egg. Furthermore, the lower chamber was not electrically isolated from the upper chamber, as the two chambers were connected by a low resistance pathway even with an egg seated firmly in the hole. The resistance across the open hole, measured with an ohmmeter, was about 10 kΩ, and increased by only 1–4 kΩ with an egg in the hole. These resistances are trivial compared with that across the egg membrane ($\geqslant 100\ \mathrm{M\Omega}$), and thus introduce no significant error into the electrical measurements. Sperm, however, did not cross from one chamber into the other during the experiments. Microscopic examination of 10 eggs inseminated in the lower chamber showed that sperm were bound only to that portion of the egg (easily identified by a bleb, or deformation in the egg surface coat) that had been in the lower chamber. (Supernumerary *Urechis* sperm do not unbind rapidly from the eggs after fertilization.)

The results of six combinations of 470 and 47 mM Na$^+$ seawater with sperm added to one or the other chamber (Fig. 2, Table 1) show that a sperm must open a localized patch of Na$^+$ channels in its vicinity. Figure 2*a–d* shows that the initial amplitudes of the fertilization potentials are dependent on [Na$^+$] at the fertilization site regardless of the [Na$^+$] bathing the rest of the egg. Figure 2*e* shows that average plateau potentials (see legend) are also predicted by the [Na$^+$] at the site of fertilization. With approximately the upper 90% (see above) of the egg

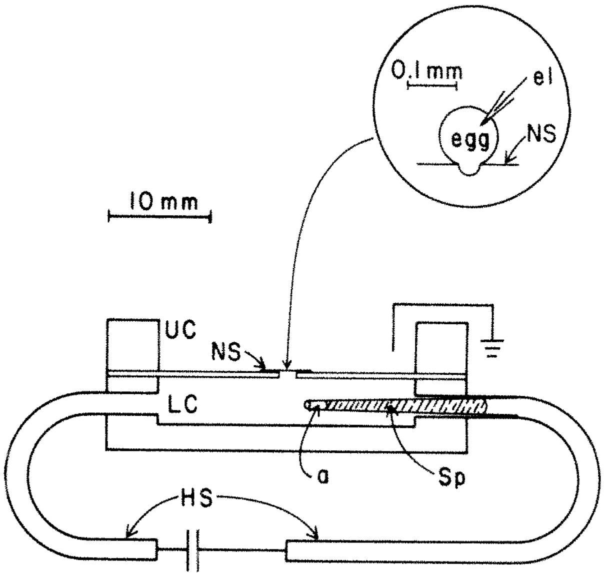


Fig. 1 Diagram of the experimental chamber modified from the original design of Robinson and Jaffe[13]. An egg (120 μm diameter) is seated in a 45-μm diameter hole in a nickel screen (NS; Perforated Products Inc.) separating two chambers (UC, upper chamber; LC, lower chamber) which can contain seawater of differing [Na$^+$]; sperm are added to one of the two chambers. A small piece of nickel screen is glued with non-toxic Paraplast over the hole in the thin plastic barrier separating upper and lower chambers. All but one hole in the nickel screen is plugged with Paraplast, and this hole is the only opening between upper and lower chambers. The respective volumes of seawater in the upper and lower chambers are 0.7 and 2 ml. Each of the two entry ports in the lower chamber is connected by thin plastic tubing to a 10-μl Hamilton syringe (HS). The left-hand syringe is used to create a slight suction that will seat the egg firmly in the hole. The right-hand syringe delivers sperm to the lower chamber. The sperm delivery tube contains about 5 μl sperm suspension (Sp) prevented from prematurely entering the chamber by an air bubble (a) at the tip. (Before inserting the tube in the chamber, external sperm are killed by washing the outside of the tube in deionized water.) To inject the sperm suspension into the lower chamber, the plungers of the two syringes are arranged back to back as shown, permitting simultaneous addition and withdrawal of equal volumes so that the position of the egg in the hole is not disturbed. Eggs were observed (at ×100 with a Wild stereomicroscope) to ensure they did not move in the hole. Sperm are delivered to the upper chamber dropwise from a pipette. At the beginning of an experiment, both upper and lower chambers are filled with seawater of the [Na$^+$] desired in the lower chamber, the sperm delivery tube is inserted into the lower chamber (or the port is plugged if sperm are to be added from above), the egg is seated in the hole and penetrated with a microelectrode, then the solution in the upper chamber is changed, if desired, by perfusion with a simple siphon plus vacuum suction system (not shown). The upper chamber was modified with plastic partitions and a baffle as described previously[7] to ensure complete solution exchange during perfusion. The solution in the upper chamber was continuously perfused whether or not its composition differed from that in the lower chamber. One-tenth [Na$^+$] seawater was prepared by diluting natural seawater with Na$^+$-free artificial seawater (NaCl replaced with choline-Cl). Methods for microelectrode measurements were those used by Jaffe *et al.*[7]. Methods for maintaining *Urechis* and obtaining and handling gametes are described elsewhere[14].

0028-0836/81/210254—03$01.00

Table 1 Average plateau potentials of the fertilization potentials obtained by varying [Na⁺] at the site of fertilization independently of [Na⁺] on the rest of the egg surface

Condition*	Average plateau potential $\bar{x}$ (mV)	s.e.m.†	(n)	P‡
$\frac{470}{470\text{s}}$	+18.2	2.4	(5)	
$\frac{47}{470\text{s}}$	+17.8	6.5	(5)	<0.02
$\frac{47\text{s}}{470}$	−14.3	4.4	(3)	>0.05
$\frac{47}{47\text{s}}$	−26.7	2.0	(3)	
$\frac{470\text{s}}{47}$	+18.0	3.0	(3)	<0.02
$\frac{470}{47\text{s}}$	+3.2	2.2	(5)	

See legend to Fig. 2*e* for method of calculation of average plateau potential. Data in the first four lines of the table are the same as in Fig. 2*e*.

*$\frac{\text{mM Na}^+\text{ in upper chamber}}{\text{mM Na}^+\text{ in lower chamber}}$, s indicates chamber to which sperm were added.

†Standard error of the mean (standard deviation/$n^{1/2}$).

‡Calculated by the *t*-test, comparing the means of the two samples indicated by the brackets.

surface in low-Na⁺ seawater and the lower 10%, to which sperm are added, in high-Na⁺ seawater, the average fertilization potential amplitude, +17 mV, is the same as that obtained with high [Na⁺] bathing the entire egg. However, if sperm are added to the upper egg surface in low [Na⁺], the average plateau potential, −14 mV, is significantly less (see Table 1). The −14 mV potential obtained with 47s/470 and the −27 mV potential with 47/47s are not significantly different at the 95% confidence level (Table 1, s indicates the chamber to which sperm were added). In reciprocal experiments, the [Na⁺] in the two chambers was reversed, with the upper 90% of the egg in high [Na⁺] and the lower 10% in low [Na⁺]. Again, fertilization in low [Na⁺] resulted in a plateau potential (+3 mV) significantly more negative than that (+18 mV) obtained with fertilization in high [Na⁺] (Table 1). However, a more positive potential was observed with 470/47s (+3 mV) than with 47/47s (−27 mV). This could be due to leakage of Na⁺ from the upper into the lower chamber through the egg's thick (>1 μm) surface coat. (A slight suction from upper to lower chamber is applied to hold the egg in the hole, and the solution bathing the lower surface of the egg is not renewed by perfusion.) Alternatively, the patch of Na⁺ channels opened by sperm may be large enough to include some of the egg surface in the opposite chamber.

Table 2 Fertilization potential characteristics of eggs seated in the experimental apparatus compared with controls resting in the bottom of a Petri dish

Natural seawater	Average plateau potential (mV)	Fertilization potential resistance (MΩ)	Unfertilized egg resistance (MΩ)
Eggs in holes	+17±2.3 (5)	146±23 (3)	
Controls	+25±1.3 (3)	135±20 (3)	548±55 (3)
1/10th [Na⁺] seawater			
Eggs in holes	−26±2.0 (3)		
Controls[7]	−21±1.6 (10)		

Average plateau potential was calculated for the first 3 min as described in Fig. 2 legend for eggs in this study; in the study of Jaffe *et al.*[7], values were calculated for the first 5 min. Resistance measurements for the controls were obtained during the 2–3-min interval after the rise of the fertilization potentials, by passing pulses of inward current ranging from −0.2 to -3×10^{-10}A, and calculating the slope resistance of the resulting current–voltage plots, which were linear. Slope resistances for the unfertilized eggs were obtained in the same way except that the membrane potential was first set to a value (+15 to +24 mV) comparable with that during a fertilization potential, and the inward current pulses superimposed. These $I-V$ plots were also linear. Resistances for the eggs in holes were determined from the voltage deflection resulting from a single inward current pulse (-1.8×10^{-10}A) at 2 min after the rise of the fertilization potential. For experiments with eggs in holes, sperm were added to the lower chamber. The values are averages ±s.e.m. (standard deviation/$n^{\frac{1}{2}}$), with n in parentheses.

Controls indicate that the observed 'patch' is not an artefact of the experimental conditions. First, the fertilization potentials of eggs in the experimental apparatus appear normal. The maximum amplitudes and average plateau potentials of eggs seated in holes are similar to those of control eggs resting in a grooved plastic dish (Table 2). In addition, measurements of egg membrane resistance during fertilization potentials in natural seawater showed that sperm-induced increases in egg membrane conductance ($G=1/R$) are also similar in experimental and control eggs; in both situations, the conductance increased approximately fourfold over that in unfertilized eggs (Table 2, columns 2, 3), indicating that the opening of Na⁺ channels is not significantly restricted by the experimental chamber. Second, eggs seated in the nickel screen activate normally. Activation was scored in four of the eggs that were fertilized from below in normal seawater. While the eggs remained seated in the hole, all four underwent normal germinal vesicle breakdown, although the germinal vesicle was in the upper portion of the egg, and in all four, surface coat elevation occurred around the entire egg. Three of these eggs were removed from the hole and cultured further; all formed polar bodies and cleaved.

The finding that sperm open a patch of Na⁺ channels is consistent with our earlier suggestion (see above and ref. 7) that the fertilization potential may be analogous to the receptor potential plus action potential system found in many chemo- and mechanosensitive cells. The Na⁺ channels are not contributed by the sperm membrane, but must pre-exist in the egg because fertilization potentials can be elicited by artificial agents (trypsin[8]) in the absence of sperm[7]. However, we do not know whether the Na⁺ channels are opened by an interaction of sperm and egg membrane components before bilayer fusion (the neurotransmitter–receptor analogy) or whether fusion must occur to initiate a fertilization potential. Whether sperm–egg fusion occurs rapidly enough (within a few seconds of insemination) to initiate the fertilization potential is not known (see Discussion, ref. 7). It also remains to be determined whether the Na⁺ channels are opened by a direct interaction of sperm and

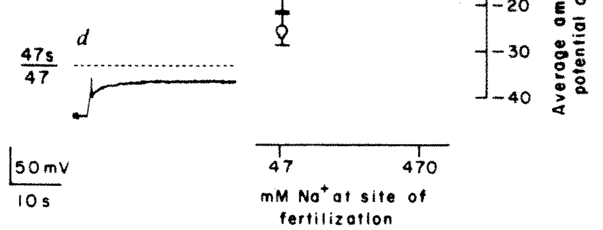


Fig. 2 Fertilization potentials obtained by varying the [Na⁺] at the site of fertilization independently of that over the rest of the egg surface. *a–d*, Original chart records showing initial portions of fertilization potentials obtained in the conditions indicated: $\frac{\text{mM Na}^+\text{ in upper chamber}}{\text{mM Na}^+\text{ in lower chamber}}$, s indicates chamber to which sperm were added. *e*, Average plateau potentials as a function of [Na⁺] at the site of fertilization. The values were obtained by averaging the potentials at 0.5, 1, 1.5, 2, 2.5 and 3 min after the initial rise of the fertilization potentials. *Urechis* egg fertilization potentials typically remain positive for about 7 min. The bars indicate the standard deviation of the mean (see Table 1 for the value of *n*). ●, Both chambers contain 470 mM Na⁺, sperm added from below; ○, the lower chamber contains 470 mM Na⁺ and the upper chamber 47 mM Na⁺; ○, both chambers contain 47 mM Na⁺, sperm added from below.

egg membrane components, or whether intermediate 'messengers' (for example, Ca^{2+}) are involved. Based on iontophoretic injection experiments, Cross[9] has proposed that a sperm-induced rise in free Ca^{2+} beneath the egg plasma membrane may initiate the fertilization potential in frog eggs. Another interesting question that remains to be answered is whether the patch of open Na^+ channels is confined to the site of sperm egg membrane contact, or whether it extends laterally.

In addition to contributing to our understanding of the gating mechanism for the fertilization potential, the results reported here raise the question of the possible significance of the 'patch' for later development. In *Urechis* eggs, the plane of the first cleavage generally passes through the point of sperm entry (within 10° of this point in 70% of the eggs[10,11]). The localized patch of open Na^+ channels may be involved in establishing this polarity[12].

This work was supported by NSF grants PCM76-84552 and PCM80-03759 to M.G-S. and by NIH grant 5 SO7RR07011-14 to the University of California, San Diego. I thank J. Raper, who was midwife to the idea for the experiment, K. Robinson for his gifts of a prototype chamber and nickel screen, and L. Holland, L. A. Jaffe and S. Hagiwara for helpful discussions.

Received 29 October 1980; accepted 4 March 1981.

1. Hagiwara, S. & Jaffe, L. A. *A. Rev. Biophys. Bioengng* **8,** 385–416 (1979).
2. Jaffe, L. A. *Nature* **261,** 68–71 (1979).
3. Gould-Somero, M., Jaffe, L. A. & Holland, L. Z. *J. Cell Biol.* **82,** 426–440 (1979).
4. Miyazaki, S. I. & Hirai, S. *Devl Biol.* **70,** 327–340 (1979).
5. Cross, N. L. & Elinson, R. P. *Devl Biol.* **75,** 187–198 (1980).
6. Nuccitelli, R. *Devl Biol.* **76,** 499–504 (1980).
7. Jaffe, L. A., Gould-Somero, M. & Holland, L. Z. *J. gen. Physiol.* **73,** 369–492 (1979).
8. Paul, M. *Devl. Biol.* **43,** 299–312 (1975).
9. Cross, N. L. *Devl Biol.* (in the press).
10. Taylor, C. V. *Physiol. Zool.* **4,** 423–460 (1931).
11. Tyler, A. *Biol. Bull.* **60,** 187–211 (1931).
12. Jaffe, L. F. in *Membrane Transduction Mechanisms* (eds Cone, R. A. & Dowling, J. E.) 199–231 (Raven, New York, 1979).
13. Robinson, K. R. & Jaffe, L. F. *Science* **187,** 70–72 (1975).
14. Gould, M. in *Methods in Developmental Biology* (eds Wilt, F. & Wessels, N.) 163–167 (Crowell, New York, 1967).

Stringency without ppGpp accumulation

**Antonella Spadaro, Angelo Spena,
Vittorio Santonastaso & Pierluigi Donini**

International Institute of Genetics and Biophysics, Via Marconi 10,
80125 Naples, Italy

Wild-type *Escherichia coli* deprived of a required amino acid shows a stringent response involving various alterations in the cell's normal physiology, including the arrest of net synthesis of stable (ribosomal and transfer) RNA, lipids, phospholipids and peptidoglycans[1]. The stringent response is found in cells of the Rel^+ (stringent) but not Rel^- (relaxed) phenotype. The intracellular concentration of guanosine 5′-diphosphate-3′-diphosphate (ppGpp) and guanosine 5′-triphosphate-3′-diphosphate (pppGpp) increases dramatically in amino acid-deprived Rel^+ cells due to the enzymatic activity of the *relA* gene product[2], which also stimulates ppGpp synthesis *in vitro*[3]. The correlation between stringent response and increase in the intracellular concentration of ppGpp has been interpreted as evidence that this nucleotide has a causal role in stringency[1]. However, as recent findings show that the correlation does not always hold, ppGpp cannot be considered to act as a unique effector molecule and the idea of a causal role for ppGpp in stringent control has become questionable[4,5]. Here we describe experiments that cast further doubt on the proposed biological role of this nucleotide, as they show that stringent control can occur in the absence of ppGpp accumulation.

The genome and the physiology of *Salmonella typhimurium* are closely related to those of *E. coli*, and it can be safely assumed that the principal aspects of stringent control do not differ substantially in the two species. In fact, in *S. typhimurium* the net synthesis of stable RNA has been shown to be under stringent control and *relA* mutants similar to those found in *E. coli* have been isolated[6]. In addition, ribosomes prepared from *relA*$^+$ *S. typhimurium* produce ppGpp *in vitro*[7]. We were therefore not surprised to find that amino acid deprivation of cultures of *relA*$^+$ *S. typhimurium* auxotrophs normally causes the accumulation of ppGpp to the same high levels observed in *E. coli* in the same conditions. However, one notable exception is shown in Fig. 1. When strain TA1001 (*hisT*, *hisD* 2655) was deprived of its sole amino acid requirement, histidine, the level of ppGpp did not rise. Even with repeated experiments there was only an occasional modest (at most twofold) increase in the ppGpp level which did not last for >5 min. When such an increase was observed, it was also seen in control experiments where the culture was filtered and resuspended in medium containing histidine. We therefore attribute this nonreproducible, slight and transient increase in ppGpp to the filtration procedure itself and not to amino acid deprivation. In these experiments the level of pppGpp never rose to values detectable by our methods.

Strain TA1001 was obtained by co-transducing the markers *aroC*$^+$, *hisT*1514 and *purF*$^+$ into strain TA997 (*aroC*5, *purF*145, *hisD*2655) (R. Cortese, personal communication). Figure 1 shows that when strain TA997 was deprived of histidine, both ppGpp and pppGpp rose to high levels, as expected for a *relA*$^+$ strain. The basal level of ppGpp was similar in both strains.

Strain TA1001 has no amino acid requirements other than histidine. It can, however, be deprived of serine by treatment with the amino acid analogue, serine hydroxamate. After such

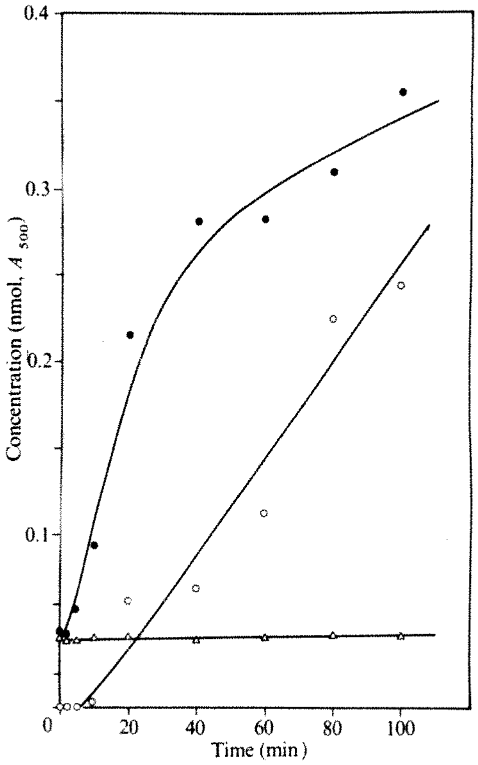


Fig. 1 Intracellular concentration of ppGpp and pppGpp in strains TA1001 and TA997 deprived of histidine. Bacteria were grown at 37 °C in low-phosphate medium[4] containing 10^{-3} M H_2HPO_4 to a concentration of 2×10^8 cells per ml, 100 μCi ml^{-1} $^{32}H_3PO_4$ were added and after one doubling the cells were filtered on Millipore filters and resuspended in medium lacking histidine, but containing 100 μ Ci ml^{-1} $^{32}H_3PO_4$. Samples were taken and analysed as described elsewhere[4]. △, ppGpp, strain TA1001; ●, ppGpp, strain TA997; ○, pppGpp, strain TA997.

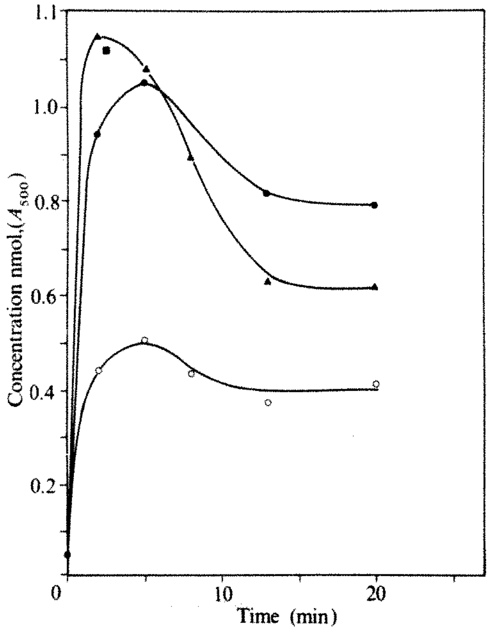


Fig. 2 Intracellular concentration of ppGpp in strain TA1001 and auxotrophic derivatives of TA1001 deprived of different amino acids. The same procedure was used as for Fig. 1. ●, Strain TA1001 deprived of serine by addition of 400 μg ml^{-1} serine hydroxamate; ○, strain PD11 deprived of threonine; ▲, strain PD12 deprived of arginine. Strains PD11 (*his*D2655, *his*T1514, *thr*A557::Tn*10*) and PD12 (*his*D2655, *his*T1514, *arg*B1837::Tn*10*) were obtained by transduction into TA1001.

treatment, very high levels of ppGpp and pppGpp are produced (see Fig. 2). Strain TA1001 can also be transduced by phage P22 carrying *S. typhimurium* DNA with the transposon Tn*10* inserted in and inactivating a gene belonging to an amino acid biosynthetic pathway. Figure 2 also shows that two such derivatives of TA1001, one requiring arginine and the other threonine, also produce ppGpp and pppGpp when deprived, respectively, of these two amino acids. Thus, strain TA1001 is not intrinsically incapable of raising its ppGpp and pppGpp levels during amino acid deprivation, but does not do so when deprived of histidine and perhaps other amino acids. Figures 1 and 2 show that in *S. typhimurium* the extent of the ppGpp response depends on the required amino acid, as in *E. coli*[4].

Figure 3*a*, *b* shows that deprivation of histidine causes the arrest of protein and stable RNA synthesis to the same extent in strains TA997 and TA1001, which suggests that strain TA1001 has inherited a functional *rel*A$^+$ gene from its TA997 parent. To prove that this arrest of stable RNA synthesis in strain TA1001 is dependent on *rel*A$^+$ function, the effect of histidine deprivation on stable RNA synthesis was determined in strain PD2, a *rel*A derivative of TA1001 constructed by transduction. Figure 3 shows that in this strain, stable RNA synthesis is under relaxed control, thus confirming that strain TA1001 harbours a *rel*A$^+$ gene and demonstrating that the arrest of stable RNA synthesis described in Fig. 3*b* is indeed attributable to stringent control. The level of ppGpp does not change when strain PD2 is deprived of histidine. Furthermore, no change in the ppGpp level is observed when serine deprivation is induced by serine hydroxamate addition—confirmation that the *rel*A allele is present in this strain.

Strains TA997 and TA1001 are isogeneic with the exception of the small segment of bacterial genome carrying the cotransducible markers *pur*F, *his*T and *aro*C. The fact that ppGpp metabolism is normal in TA997 strongly suggests that the defect present in TA1001 harbours one or more additional mutations close to *his*T. The *his*T gene product catalyses the pseudouridine modification in the anticodon region of the many tRNAs that contain pseudouridine in this region[8]. Mutations in the *his*T gene alter the regulation of several amino acid biosynthetic pathways[8], and histidine tRNA derived from a *his*T mutant is defective in translation[9]. Thus, it is possible that the alteration in ppGpp and pppGpp metabolism in strain TA1001 is another consequence of the subnormal pseudouridylation of histidine tRNA. The three amino acids which produce a normal response (Fig. 2) in TA1001 are cognate to tRNA species all or most of which have no pseudouridine in the anticodon region, and are therefore not affected by the *his*T mutation. Note that one of a different class of histidine regulatory mutants, *his*U, is altered in the control of the accumulation of stable RNA during carbon and energy source transitions[10].

We do not know whether the abnormal response of TA1001 to histidine deprivation is due to a block in the production of

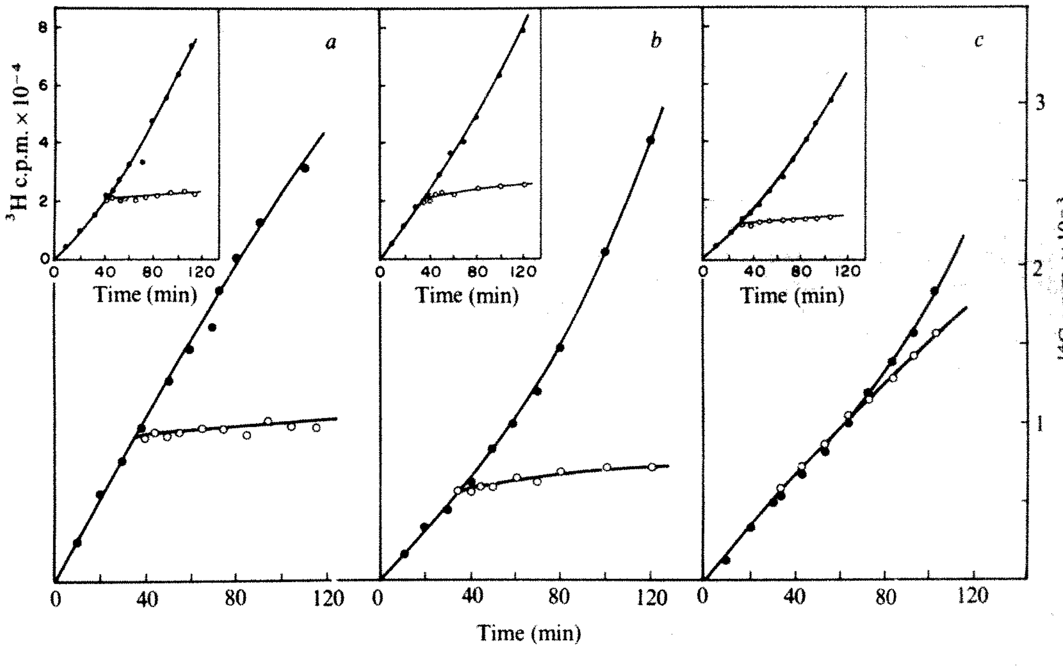


Fig. 3 RNA and protein synthesis in strains TA997, TA1001 and PD2 deprived of histidine. Bacteria were grown to a concentration of 4 × 10^8 cells per ml in low-phosphate medium, and ^{14}C-uridine (0.01 μ Ci ml^{-1}, 5 μg ml^{-1}) and ^{3}H-isoleucine (1 μ Ci ml^{-1}, 3.5 μg ml^{-1}) were added. Half the culture was filtered after 30 min and resuspended in medium lacking histidine but containing the labelled precursors at the same specific activities as before. Samples were taken up in 5% TCA, filtered on Millipore filters and counted in a Packard Tri-Carb scintillation counter. Large graphs show incorporation into RNA; insets show incorporation into protein. *a*, TA997; *b*, TA1001; *c*, PD2. ●, Not deprived; ○, deprived. PD2 (*his*D2655, *his*T1514, *rel*A1::Tn*10*) was constructed by transducing *rel*A1::Tn*10* into TA1001.

ppGpp or to an increase in its degradation. The main conclusion that can be drawn from the above observations is that an increase in the intracellular concentration of ppGpp is not a necessary condition for stringent control of the net synthesis of stable RNA. This leads to one of the two following conclusions: either ppGpp has no causal role in stringent control of RNA accumulation, or inhibition by ppGpp is not the only mechanism, but exists alongside and competes with at least one other mechanism also dependent on the *rel*A$^+$ gene. In case of failure of the ppGpp mechanism, the putative accessory mechanisms would become totally responsible for stringent control. With respect to the control of RNA accumulation, the latter hypothesis finds some support in several studies which have shown that a considerable fraction of the decrease in net synthesis of stable RNA can be attributed to degradation of the RNA[11,12]. Thus, it may be that in TA1001, stringent control of RNA synthesis is totally degradational.

The results obtained with strain TA1001 also make it rather unlikely that the effector molecule in stringent control could be a derivative of ppGpp, such as guanosine 5′-diphosphate-3′-monophosphate (ppGp), as has been suggested[5]. The role of the guanosine polyphosphates in bacterial physiology is now, more than ever, an open question which a thorough examination of the properties of strain TA1001 may elucidate.

We thank Drs Giovanna Ames, John Roth, Riccardo Cortese and Francesco Blasi for providing strains and helpful advice.

Received 9 December 1980; accepted 4 March 1981.

1. Gallant, J. A. *A. Rev. Genet.* **13,** 393–415 (1979).
2. Cashel, M. *J. biol. Chem.* **244,** 3133–3141 (1969).
3. Haseltine, W. A., Block, R., Gilbert, W. & Weber, K. *Nature* **238,** 381–384 (1972).
4. Donini, P., Santonastaso, V., Roche, J. & Cozzone, A. J. *Molec. Biol. Rep.* **4,** 15–29 (1978).
5. Pao, C. C. & Gallant, J. *J. biol. Chem.* **254,** 688–692 (1979).
6. Martin, R. G. *J. mol. Biol.* **31,** 127–134 (1968).
7. Stephens, J. C., Artz, S. W. & Ames, B. N. *Proc. natn. Acad. Sci. U.S.A.* **72,** 4389–4393 (1975).
8. Turnbough, C. L. Jr., Neill, R. J., Landsberg, R. & Ames, B. N. *J. biol. Chem.* **254,** 5111–5119 (1979).
9. Johnston, M. M., Barnes, W. M., Forrest, G. C., Bossi, L. & Roth, J. R. *Proc. natn. Acad. Sci. U.S.A.* **77,** 508–512 (1980).
10. Davidon, J. P. & Williams, L. S. *Biochem. biophys. Res. Commun.* **88,** 682–687 (1979).
11. Donini, P. *J. molec. Biol.* **72,** 553–569 (1972).
12. Pedersen, S. in *Control of Ribosome Synthesis* (eds Kjeldgaard, N. C. & Maaloe, O.) 345–352 (Academic, New York, 1976).

Nuclear sodium and potassium

Philip L. Paine, Terry W. Pearson, Louis J. M. Tluczek & Samuel B. Horowitz

Cellular Physiology Laboratory, Department of Biology, Michigan
Cancer Foundation, 110 East Warren Avenue, Detroit,
Michigan 48201, USA

Na$^+$ and K$^+$ influence gene expression[1–3], and it is important to know the concentrations and chemical activities of these cations in the cell nucleus and to understand their control. However, measurements are hampered by the small size of nuclei, and also because (1) the steep Na$^+$ and K$^+$ activity gradients across cell membranes can cause experimental manipulations to provoke rapid (~0.1 s half-time for a 10 μm cell) artefactual cation redistributions; (2) intracellular activities cannot be directly inferred from concentrations because Na$^+$, K$^+$ and water exist partly in bound form. We circumvent these difficulties by using a giant cell (amphibian oocyte), cryogenic methods to limit diffusion, and an artificial intracellular reference phase[4,5]. Century *et al.*[6] found K$^+$ more concentrated in the oocyte nucleus than in cytoplasm, and the reverse for Na$^+$. Our data for the *Desmognathus o. ochrophaeus* (salamander) oocyte are similar: the nucleus contains 30±3 μequiv Na$^+$ and 144±3 μequiv K$^+$ per ml H_2O; the cytoplasm contains 87±2 μequiv Na$^+$ and 87±2 μequiv K$^+$ per ml H_2O (means±s.e.m.). These data could imply that the nucleus actively transports Na$^+$ and K$^+$, but strong evidence suggests otherwise: the nuclear envelope, with pores of effective radius 45 Å (ref. 7), is too permeable to maintain small solute concentration gradients[8–10], and microelectrodes measure equal nuclear and cytoplasmic cation activities[11–13]. Here, we confirm and extend these findings, showing that nucleus/cytoplasm cation concentration differences are due to (1) partial exclusion of diffusive Na$^+$ and K$^+$ by cytoplasm and (2) differential Na$^+$ and K$^+$ binding by nucleus and cytoplasm.

Oocytes were microinjected at 34 °C with 50–100 nl of 14% gelatin sol, sometimes containing ^{3}H-sucrose. The gelatin, which displaces a volume of cytoplasm, was gelled by immersing the oocyte in iced Ringer's solution for 5 min. Subsequent incubation was at 5 °C. The gelled gelatin or reference phase (RP) excludes organelles and structural elements, and its water comes to diffusional equilibrium with intracellular solutes[4,5]. Oocytes were then frozen at −160 °C in dichlorodifluoromethane, and stored in liquid nitrogen. Nucleus, RP and cytoplasm were isolated at −45 °C by ultralow-temperature microdissection, weighed, dried and reweighed to determine water content, and extracted in water by boiling[6,14]. Cation concentrations were determined by atomic absorption spectroscopy and ^{3}H-sucrose by liquid scintillation counting.

Intracellular ^{3}H-sucrose does not distribute uniformly at equilibrium[5,15], nuclear and RP concentrations being equal, but about three times that of the cytoplasm (Fig. 1). Because sucrose distributes uniformly when gelatin is dialysed against saline[15], we deduce that the nucleus/cytoplasm asymmetry is not due to nuclear binding, but to sucrose's reduced solubility in, or exclusion from, cytoplasmic water.

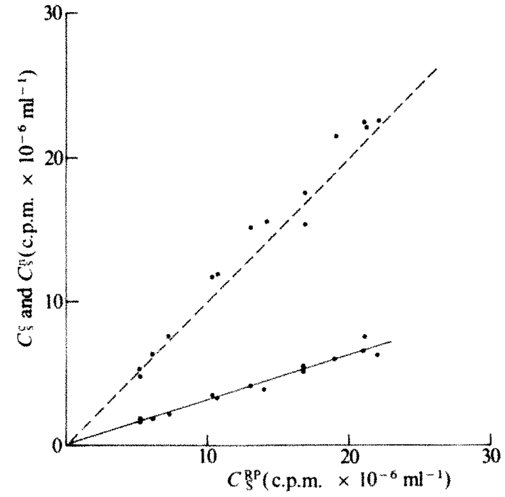


Fig. 1 Nucleus/RP (dashed line) and cytoplasm/RP (solid line) represent concentrations of ^{3}H-sucrose (C_s). Oocytes were manually dissected from ovaries in ice-cold Ringer's solution containing (mM) 104.5 NaCl, 2.5 KCl, 1.2 $MgSO_4$, 6.6 $NaHCO_3$, 1.2 Na_2HPO_4, 2.0 NaH_2PO_4 and 0.7 $CaCl_2$ (*p*H 7.2), and injected with gelatin containing ^{3}H-sucrose, incubated for 4–20 h, dissected by ultralow-temperature microdissection, and analysed as described in the text. C_s^c, C_s^n and C_s^{RP} are ^{3}H-sucrose concentrations (per ml H_2O) in cytoplasm, nucleus and RP, respectively. The least-squares line for cytoplasm/RP data is $C_s^c = 0.31C_s^{RP} + 0.01$ (×10^6 c.p.m. ml^{-1}). The dashed line is isomolar, $C_s^n = C_s^{RP}$; the least-squares line determined for the nucleus/RP data (not shown) is $C_s^n = 1.04C_s^{RP} + 0.15$ (×10^6 c.p.m. ml^{-1}).

0028-0836/81/210258—04$01.00

Figure 2 compares Na$^+$ and K$^+$ concentrations, C_{Na} and C_K, in RP, nucleus (n) and cytoplasm (c). C_{Na}^{RP} and C_K^{RP} are reciprocally related, their sum constant at ~151 μequiv ml^{-1} (Fig. 2*a*). There is a clustering of cells with $C_K^{RP}/(C_{Na}^{RP}+C_K^{RP})>0.7$ (Fig. 2, inset). We consider these cells 'normal' because their cytoplasmic and nuclear cation concentrations resemble those of uninjected cells[5]. Typical of eukaryotic cells, Na$^+$ and K$^+$ activities (measured by the RP) in normal oocytes are, respectively, less and greater than those of the Ringer's (Fig. 2 legend). $C_K^{RP}/(C_{Na}^{RP}+C_K^{RP})<0.7$ results from damage during microinjection and experimental manipulation, but these values allow isothermal analysis of mechanisms. Nuclear data resemble RP data, but most nuclear points fall above the RP line (Fig. 2*b*). (In normal oocytes $C_{Na}^n+C_K^n=171\pm4$ μequiv ml^{-1}.) Also, when $C_{Na}^n>100$ μequiv ml^{-1}, Na$^+$ increase is not stoichiometrically compensated by K$^+$ loss, as in the RP. C_K^c and C_{Na}^c are also reciprocally related, but C_K^c is always <105 μequiv ml^{-1} and $C_{Na}^c>70$ μequiv ml^{-1} (Fig. 2*c*). Obviously, cytoplasm, RP and nucleus handle cations differently.

Cytoplasmic and RP Na$^+$ are not isomolar (Fig. 3*a*, open circles). As gelatin/saline dialysis isotherms are isomolar[4], these data demonstrate the existence of two cytoplasmic Na$^+$ fractions, one shared with the RP and therefore diffusive, and the other restricted to the cytoplasm, or bound. The concentrations of diffusive and bound Na$^+$ (relative to total cytoplasmic water) are denoted by C_{Na}^{c1} and C_{Na}^{c2}, respectively, so that $C_{Na}^c=C_{Na}^{c1}+C_{Na}^{c2}$. Like sucrose, diffusive Na$^+$ has reduced solubility in cytoplasm, the cytoplasm/RP diffusive Na$^+$ distribution coefficient (q_{Na}) being 0.45 (the slope in Fig. 3*a*). The intercept, 75 μequiv ml^{-1}, is the saturation concentration of bound cytoplasmic Na$^+$, max-C_{Na}^{c2}.

The nucleus behaves similarly to the RP (Fig. 3*b*). At low Na$^+$ concentrations nuclear and RP data are superimposable, showing that nucleus and cytoplasm share only diffusive Na$^+$ and that the nucleus interacts with Na$^+$ as though it were an RP or ordinary saline solution (Fig. 3*b*). When $C_{Na}^n>100$ μequiv ml^{-1}, increase in RP (diffusive) Na$^+$ is accompanied by a larger increase in nuclear concentration. Perhaps the nucleus contains macromolecular charge groups which are unmasked when exposed to Na$^+>100$ μequiv ml^{-1} (refs 8, 16).

Cytoplasmic K$^+$, like Na$^+$, consists of two fractions, a diffusive fraction, of concentration C_K^{c1}, shared with the RP, and an unshared, nondiffusive fraction, of concentration C_K^{c2} (Fig. 4*a*). The cytoplasm/RP distribution coefficient for diffusive K$^+$, q_K, is 0.36 (slope, Fig. 4*a*). The intercept at 41 μequiv ml^{-1} is the saturation concentration of bound cytoplasmic K$^+$, max-C_K^{c2}. RP and nucleus concentrations are similar, differing only in a small (19 μequiv K$^+$ ml^{-1}) invariant nuclear excess (Fig. 4*b*). Apparently, nuclear water acts like ordinary or RP water as solvent for K$^+$, but in addition the nucleus contains 19 μequiv ml^{-1} of bound K$^+$.

A previous study of *Rana pipiens* oocytes incubated in ^{42}K$^+$-Ringer's[17] revealed two cytoplasmic K$^+$ fractions indistinguishable from those demonstrated here. Exchange of the fast isotopic fraction, present in both nucleus and cytoplasm, was limited by cell-membrane permeability. The second fraction, restricted to cytoplasm, exchanged too slowly to be measured. As short-term incubations suggested that ^{42}K$^+$ reflects the distribution of diffusive K$^+$, we tested this by incubating RP oocytes for 3–4 h in ^{42}K$^+$-Ringer's, and found that the cytoplasm/RP concentration ratio of ^{42}K$^+$, q_K^*, was 0.36, identical to q_K from the K$^+$ isotherm. Hence, ^{42}K$^+$ exchange with cytoplasmic bound K$^+$ is negligible, and diffusive K$^+$, like sucrose and Na$^+$, is more concentrated in the RP than in cytoplasm. Also, the ^{42}K$^+$ nucleus/RP ratio for normal oocytes was 1.14 ± 0.03, agreeing with Fig. 4*b* in demonstrating nondiffusive nuclear K$^+$. However, bound nuclear K$^+$ exchanges completely with ^{42}K$^+$ (nucleus/RP specific activities ratio $=1.00\pm0.03$) in <3 h, while cytoplasmic exchange is extremely slow.

At equilibrium, the RP—an ordinary water phase with no membrane—and the nucleus are essentially isomolar for diffusive Na$^+$, K$^+$ and sucrose, demonstrating that the nucleus

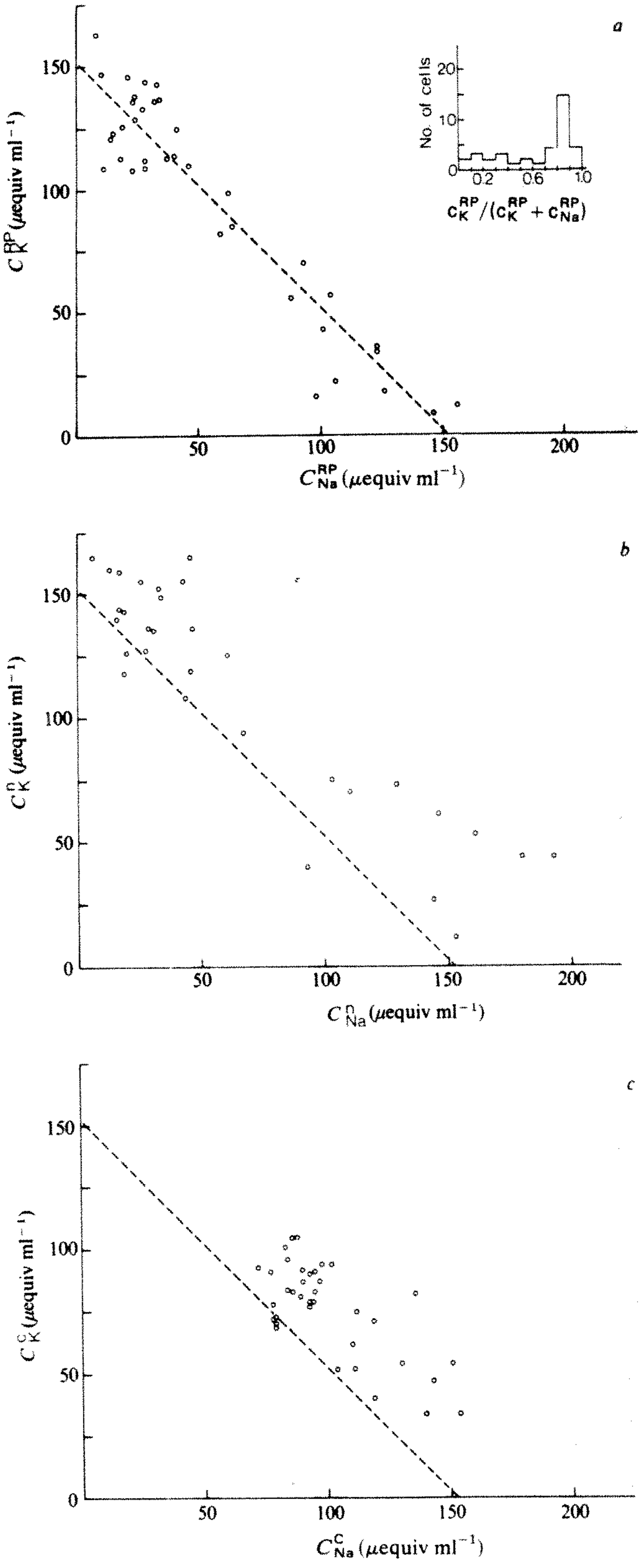


Fig. 2 Concentrations of K$^+$ and Na$^+$ (C_K and C_{Na}) in RP, nucleus and cytoplasm. *a*, RP data for cells fitted by the least-squares line, $C_K^{RP}=-C_{Na}^{RP}+151$ μequiv ml^{-1}. (Inset, frequency distribution of cells as a function of fractional RP K$^+$ content.) Equilibrium dialysis shows that RP gelatin has ordinary water solvent properties. Hence, intracellular solute activities (a_i) can be determined from $a_i=C_i^{RP}\gamma_i^0$, in which $\gamma_i^0\sim0.77$ is the activity coefficient in salines of comparable concentrations[5]. In normal oocytes ($C_K^{RP}/(C_{Na}^{RP}+C_K^{RP})>0.7$) cytoplasmic and nuclear activities are 99 μequiv K$^+$ ml^{-1} and 18 μequiv Na$^+$ ml^{-1}, compared with Ringer's 2 μequiv K$^+$ ml^{-1} and 89 μequiv Na$^+$ ml^{-1}. *b*, *c*, Data for nucleus and cytoplasm, respectively. Least-squares line from RP data (dashed) is shown for comparison.

does not actively transport cations or sucrose. Furthermore, the RP directly reflects effective cytoplasmic free concentrations and hence chemical activities[4]. Except for a small amount of binding, the same is true for the nucleus, which also behaves toward microsolutes like an ordinary aqueous phase embedded in a complex cytoplasm. Nuclear solute control can therefore only be understood when we know exactly how cytoplasm differs from simple aqueous solutions. Phenomenologically, two factors are identifiable: exclusion and binding. Cytoplasmic exclusion of solute i is indexed by q_i, the cytoplasm/nucleus or cytoplasm/RP equilibrium distribution coefficient of the diffusive form. If binding is negligible, as with sucrose and $^{42}K^+$ at short exchange times, the coefficient is equal to C^c_i/C^n_i or C^c_i/C^{RP}_i. When binding is significant, q_i is the slope of the cytoplasm/RP isotherm asymptote. Present determinations are q_s (sucrose) = 0.31 (Fig. 1), $q^*_K = 0.36$, $q_K = 0.36$ (Fig. 4*a*), and $q_{Na} = 0.45$ (Fig. 3*a*). Bound cytoplasmic concentrations are also determined from cytoplasm/RP isotherms; at saturation, C^{c2}_{Na} and C^{c2}_K are 75 and 41 µequiv ml^{-1}, respectively. (There is evidence that when C^{RP}_i is very low, the isotherms are nonlinear[5], implying that some bound cations are in mass action equilibrium with diffusive cations.) Changes in the affinities and/or capacities of cytoplasmic binding and changes in cytoplasmic solvent capacity must be considered as possible mechanisms for transient modulation of the baseline level of nuclear solute activities maintained by the cell membrane[18].

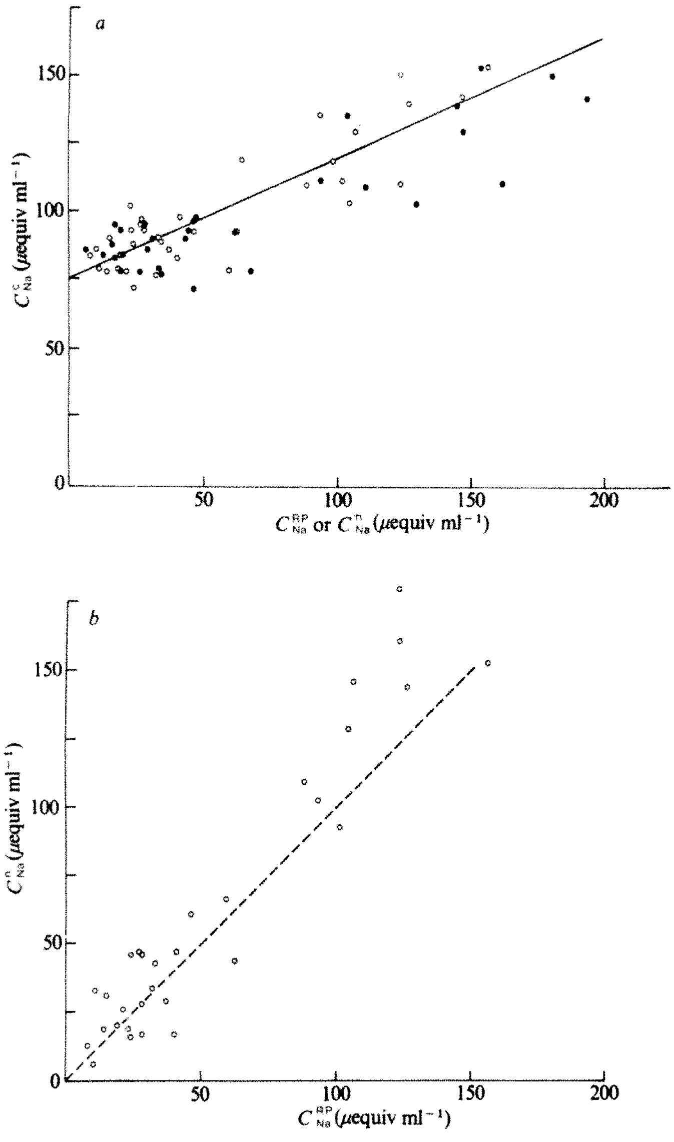


Fig. 3 Sodium isothermal relations. *a*, Cytoplasmic Na$^+$ concentration, C^c_{Na}, versus RP Na$^+$ concentration, C^{RP}_{Na} (open circles), and nuclear Na$^+$ concentration, C^n_{Na} (closed circles). The line is the least-squares line for the cytoplasm/RP data, $C^c_{Na} = 0.45\ C^{RP}_{Na} +$ 75 µequiv ml^{-1}. *b*, C^n_{Na} versus C^{RP}_{Na}. The dashed line is isomolar, the least-squares line (not shown) is $C^n_{Na} = C^{RP}_{Na} + 4$ µequiv ml^{-1}.

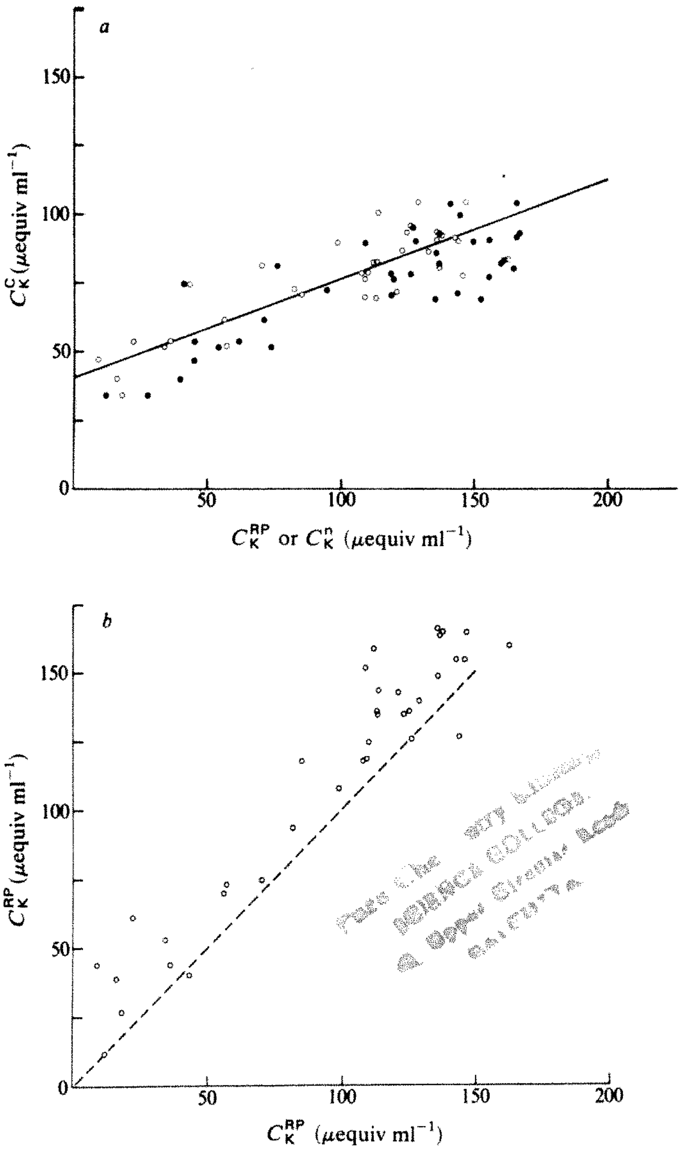


Fig. 4 Potassium isothermal relations. *a*, Cytoplasmic K$^+$ concentration, C^c_K, versus the RP K$^+$ concentration, C^{RP}_K (open circles); the least-squares line is $C^c_K = 0.36\,C^{RP}_K + 41$ µequiv ml^{-1}. C^c_K compared with nuclear K$^+$ concentration, C^n_K (closed circles); the least-squares line (not shown) is $C^c_K = 0.34\,C^n_K + 38$ µequiv ml^{-1}. *b*, C^n_K versus C^{RP}_K. The dashed line is isomolar; the least-squares line (not shown) fitting the data points is $C^n_K = C^{RP}_K + 19$ µequiv ml^{-1}.

Cytoplasmic exclusion and binding can be approached through two paradigms. One assumes that most cell water is saline-like, but is partially in organelles or vesicles, surrounded by membranes with restricted permeabilities and/or active transport properties[19]. The second hypothesizes decreased solubility due to structuring of cell water[5,18]. Vesicle and structured water mechanisms are not mutually exclusive. Oocyte cytoplasm contains abundant candidates (ground cytoplasm, endoplasmic reticulum, lysozomes, mitochondria, yolk platelets) to account for exclusion and binding within either or both paradigms.

Exclusion and binding seem to be widely distributed properties of cells. The low nucleus (or RP)/cytoplasm Na^+ ratio which indicates that the diffusive/total Na^+ ratio in oocyte cytoplasm is low, is paralleled by Na^+-specific electrode determinations of low activity/concentration ratios ('apparent' activity coefficients) not only in oocytes[20,21], but also in neurones[22], epithelia[12,23] and muscle[22,24–26]. Similarly, the high nucleus (or RP)/cytoplasm K^+ ratio of oocytes is paralleled by high activity/concentration ratios determined with K^+ electrodes in oocytes[21] and other cells[5,27]. Deviation of cytoplasmic activity coefficients from those of ordinary salines is direct evidence of solute binding or exclusion. However, because binding and exclusion have opposing influences on activity/concentration ratios, intracellular activity coefficients resembling those in salines do not prove that cytoplasm is a simple water phase[27]. Similarly, one cannot conclude that nucleus and cytoplasm are necessarily physicochemically similar if their Na^+ or K^+ concentration ratios are close to unity.

We thank Drs P. Furmanski, Y.-T. Lau and J. K. Reynhout for helpful criticism. This work was supported by NIH grants GM 19548, GM 26734 and HD 12512, and by an institutional grant from the United Foundation of Greater Detroit.

Received 15 September 1980; accepted 17 February 1981.

1. Lezzi, M. & Gilbert, L. I. *J. Cell Sci.* **6,** 615–627 (1970).
2. Rensing, L. & Fisher, M. *Cell Differentiation* **4,** 209–217 (1975).
3. Kroeger, H., Trösch, W. & Müller, G. *Expl Cell Res.* **80,** 329–339 (1973).
4. Horowitz, S. B., Paine, P. L., Tluczek, L. & Reynhout, J. K. *Biophys. J.* **25,** 33–44 (1979).
5. Horowitz, S. B. & Paine, P. L. *Biophys. J.* **25,** 45–62 (1979).
6. Century, T. J., Fenichel, I. R. & Horowitz, S. B. *J. Cell Sci.* **7,** 5–13 (1970).
7. Paine, P. L., Moore, L. C. & Horowitz, S. B. *Nature* **254,** 109–114 (1975).
8. Paine, P. L. & Horowitz, S. B. in *Cell Biology: A Comprehensive Treatise* Vol. 4 (eds Goldstein, L. & Prescott, D. M.) 299–338 (Academic, New York, 1980).
9. Harris, J. R. *Biochim. biophys. Acta* **515,** 55–104 (1978).
10. Siebert, G. *Biochem. Soc. Trans.* **6,** 5–9 (1978).
11. Starodubov, S. M. & Kurella, G. A. *J. gen. Bot.* **33,** 464–468 (1972).
12. Palmer, L. G. & Civan, M. M. *J. Membrane Biol.* **33,** 41–61 (1977).
13. Wuhrmann, P., Ineichen, H., Riesen-Willi, U. & Lezzi, M. *Proc. natn. Acad. Sci. U.S.A.* **76,** 806–808 (1979).
14. Frank, M. & Horowitz, S. B. *J. Cell Sci.* **19,** 127–139 (1975).
15. Horowitz, S. B. & Paine, P. L. *Nature* **260,** 151–153 (1976).
16. Itoh, S. & Schwartz, I. L. *Am. J. Physiol.* **188,** 490–498 (1957).
17. Frank, M. & Horowitz, S. B. *Am. J. Physiol.* **238,** C133–C138 (1980).
18. Ling, G. N., Miller, C. & Ochsenfeld, M. M. *Ann. N.Y. Acad. Sci.* **204,** 6–50 (1973).
19. Dick, D. A. T. *J. Physiol., Lond.* **284,** 37–53 (1978).
20. Dick, D. A. T. & McLaughlin, S. G. A. *J. Physiol., Lond.* **205,** 61–78 (1969).
21. Palmer, L. G., Century, T. J. & Civan, M. M. *J. Membrane Biol.* **40,** 25–38 (1978).
22. Kostyuk, P. G., Sorokina, Z. A. & Kholodova, Yu. D. in *Glass Microelectrodes* (eds Lavellee, M., Schanne, O. F. & Hebert, N. C.) 322–348 (Wiley, New York, 1969).
23. Lee, C. O. & Armstrong, W. McD. *Science* **175,** 1261–1264 (1972).
24. Lev, A. A. *Nature* **201,** 1132–1134 (1964).
25. Armstrong, W. McD. & Lee, C. O. *Science* **171,** 413–415 (1971).
26. Lee, C. O. & Armstrong, W. McD. *J. Membrane Biol.* **15,** 331–362 (1974).
27. Lev, A. A. & Armstrong, W. McD. *Curr. Topics Membrane Transport* **6,** 59–123 (1975).

Structural homologies between prealbumin, gastrointestinal prohormones and other proteins

Hans Jörnvall[*], Anders Carlström[†], Tom Pettersson[†], Björn Jacobsson[‡], Magnus Persson[§] & Viktor Mutt[¶]

Departments of Chemistry I[*] and Biochemistry II[¶], Karolinska Institutet, S-104 01 Stockholm 60, Departments of Clinical Chemistry[†] and Pathology[‡], Karolinska Institutet, Danderyds Hospital, S-182 88 Danderyd, and Department of Numerical Analysis[§], Royal Institute of Technology, S-100 44 Stockholm 70, Sweden

Prealbumin has unusual properties for a plasma protein. It is an unglycosylated small protein composed of four identical 127-residue polypeptide chains, and although generally thought to be of hepatic origin[1], its site of synthesis has not been shown. However, its amino acid sequence[2–4], conformation and molecular interactions[5–8] are known. Prealbumin has a role in retinol[9] and thyroxine[10] transport, and in the metabolic regulation mediated by thyroid hormones[11,12]. A thymic hormone-like activity has also been reported[13]. Prealbumin-like immunoreactivity has been demonstrated in the α cells of the pancreas[14] and in endocrine cells of the small intestine (T.P., B.J., B. Sandstedt and A.C., unpublished results). A relationship with some activity of the gastrointestinal tract is therefore possible. Here we show sequence similarities between prealbumin and the glucagon–secretin type of gastrointestinal hormones, as well as between prealbumin and retinol-binding protein. Possible evolutionary links with prohormones and other functions in a new protein family are demonstrated—these may explain the unusual properties of prealbumin.

The primary structures of mammalian glucagon, secretin, vasoactive intestinal peptide (summarized in ref. 15), gastric inhibitory peptide[16] and the C-terminal part of the possible glucagon prohormone[17], glicentin[18], were compared with the structure of human plasma prealbumin. The hormones are related, forming a family[15]. All possible hormone alignments along prealbumin were tested using a computer; randomly generated sequences with identical total compositions were similarly compared. By this method, the probability of identities shared between the actual structures belonging to the distribution of background similarities could be calculated.

With the exception of gastric inhibitory peptide, all these hormones give only one alignment with prealbumin at a low probability level, and this alignment is the same for the four hormones, starting at position 90 in prealbumin (see Fig. 1). The probabilities, single fits and identical positions suggest that the hormones align uniquely with the C-terminal part of prealbumin. Gastric inhibitory peptide (Fig. 1) falls outside this pattern, but it is very different from other members of the hormone group[15].

To test the significance of the similarities further, comparisons were made with retinol-binding protein[19] (functionally related to prealbumin) and with horse liver alcohol dehydrogenase (no known relationship to any of these polypeptides). In the case of retinol-binding protein, all sequence identities to the hormones were insignificant (fivefold greater chance occurrence than with prealbumin, Fig. 1), and alignments were not unique but multiple (for three hormones) or starting at different positions (two hormones). For alcohol dehydrogenase, probability values were variable, and the alignments were again multiple (three hormones) or single but different (two hormones). Thus, the hormone comparisons with retinol-binding protein and alcohol dehydrogenase support the significance of the similarities in Fig. 1.

The C-terminal location of the similarity as well as the size of prealbumin agree with the idea of an homology between prealbumin and a glucagon prohormone. Glicentin[18] has the glucagon sequence in the C-terminal region and a size similar to that of the prealbumin subunit[18]. The same features apply to the somatostatin precursor[20].

These independent protein data support the patterns of sequence similarities in Fig. 1. The prealbumin homology to a prohormone of the glucagon family is unexpected and resembles other super-families[15] detected by sequence comparisons, for example, lysozyme–lactalbumin[21], haptoglobin–serine proteases[22] or ovalbumin–antiproteases[23]. It is even possible that prealbumin has retained a prohormone function in addition to its known interactions with retinol-binding protein and thyroxine. It exhibits thymic hormone-like activity[13] but independent of this or other activities, the structural homology with

	90 … 100 … 110 … 120 … 127	Identities with prealbumin	Chance probability	Unique alignment
PRE	**H** **A** **E** V **V** **F** **T** A **N** D **S** G P **R** R Y T I **A** A L **L** S P Y S Y S **T** T A V V T **N** P **K** E			
SECR	**H** S D G T **F** **T** S E L **S** R L **R** D S A R L Q R **L** L Q G L V	6	0.015	+
VIP	**H** S D A **V** **F** **T** D **N** Y T R L **R** K Q M A V K K Y L N S I L N	6	0.010	+
GLIC	**H** S Q G T **F** **T** S D Y **S** K Y L D S R R **A** Q D F V Q W L M N **T** K R N K N **N** I A	7	0.013	+
GLUC	**H** S Q G T **F** **T** S D Y **S** K Y L D S R R **A** Q D F V Q W L M N **T**	6	0.010	+
GIP	Y **A** **E** G T **F** I S D Y **S** I A M D K I R Q Q D F V N W L L A Q K G K K S D W **K** H N I T Q	5	0.134	–

Fig. 1 Results of sequence comparisons between human prealbumin (PRE), pig secretin (SECR), pig and bovine vasoactive intestinal peptide (VIP), part of pig glicentin (GLIC), pig and human glucagon (GLUC) and pig gastric inhibitory peptide (GIP). Amino acid residues are shown by one-letter codes, with similarities between prealbumin and any other peptide in bold type. Positional numbers refer to prealbumin. Probability values represent the chance occurrence of alignment of each peptide independent of the others, as determined by the program described. Unique alignment indicates that no other fit of the peptides along prealbumin gives the same number of identities.

gastrointestinal hormones explains the tissue localization of prealbumin.

The structures of prealbumin and retinol-binding protein[19] were compared using all possible combinations of 30-residue segments. Three pairs of similarities (well below the 0.01 random probability level) were distinguished from all other pairs (above the 0.01 level), and could also be correlated with each other (see Fig. 2). One pair suggests an internal repetition in retinol-binding protein, and the other two pairs relate each of these two repetitions to one segment of prealbumin. The positional alignments of the three pairs detected are similar for the mutually independent comparisons, only deviating a little in the case of the prealbumin segment.

The repetitive region in retinol-binding protein involves 11 identities in 43 residues (26% identity; $P=0.0005$ chance occurrence, Fig. 2), giving values as distinctive as those for the established protein connections between yeast and horse alcohol dehydrogenases. Furthermore, the two segments in retinol-binding protein detected from internal comparisons (positions 17–59 and 58–100, Fig. 2) agree almost exactly with those (positions 15–58 and 60–97) revealed by the comparison with prealbumin. The similarities to prealbumin involve a shift in positional numbers, but are still at the 0.005 probability level (Fig. 2). If not merely reflecting conformational similarities, such shifted numbers could indicate gaps, as shown for another peptide comparison (H.J., V.M. and M.P., unpublished results).

The relative positions of the segments are summarized in Fig. 3*a*. The region marked 'common' in prealbumin has similarities that are less significant than those in the other regions. However, all common regions are compatible with an ancestral connection between prealbumin and retinol-binding protein, and a partial duplication in the latter. If significant, these regions correspond to the most conserved and common function of both proteins. One such region may be the area involved in the binding interactions between prealbumin and retinol-binding protein.

An ancestral connection between the two proteins is also supported by the size of the subunits. A duplication of the common region (consisting of ~40 residues) in retinol-binding protein, and some shifts due to gaps, make the corresponding ancestors similar in length. The present size difference is 55 residues, with prealbumin the shorter subunit, as anticipated by such a scheme.

The position of the two segments in the conformation of the prealbumin subunit is shown in Fig. 3*b*. Both regions cover large areas and cannot be ascribed to single functions. However, in the tetramer, parts of the 'hormone' regions limit the thyroxine binding site (Fig. 3*b*). The association between a primary structure homologous to one hormone and a tertiary structure binding another would, if ancestrally significant, support the idea that receptors and their ligands have a common evolutionary origin[24–26]. In the divergence from the ancestor, prealbumin would then have become a member of the receptor line rather than of the prohormone line. An ancestral connection between nuclear receptors and prealbumin has also been inferred[7] from the presence of a DNA binding site in prealbumin. The start of the hormone region (positions 90–96) forms the bottom of the tetramer's cylindrical impression which constitutes the DNA binding site (Fig. 3*b*).

The central parts of the common region in the prealbumin chain are superficial in the tetramer (Fig. 3*b*), and could therefore be involved in the binding of the retinol-binding protein. However, it is not known whether this is true, and it will be of interest to see if the sequence similarities between the two proteins correspond to folding similarities. If so, this part of retinol-binding protein would contain six β strands, as each common region in prealbumin (Fig. 2) has three β strands[8]. Anyway, known conformational arrangements in prealbumin fit the scheme suggested from the observed homologies. Opposite arrangements of the hormone and common segments would not have provided similar support.

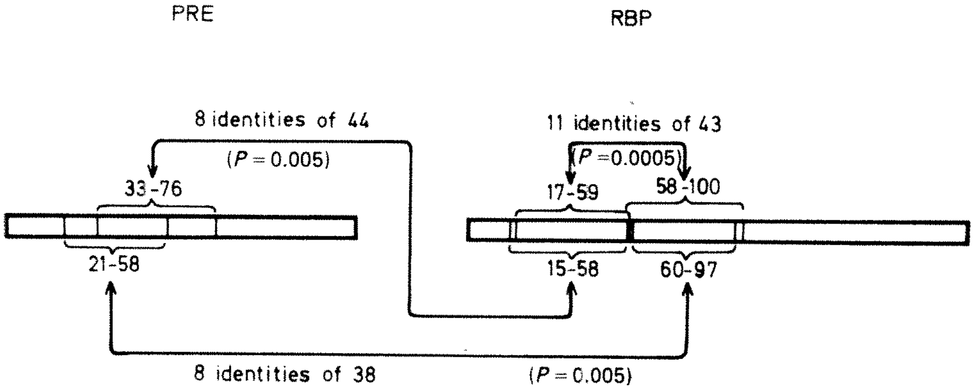


Fig. 2 Schematic representation of the three single pairs (double arrows) of sequence similarities detected in all possible combinations of segment comparisons between and within the subunits of prealbumin (PRE) and retinol-binding protein (RBP). Polypeptides (bars) and similarities (with end positions indicated) are drawn to scale. Probability values (P) represent the chance occurrence.

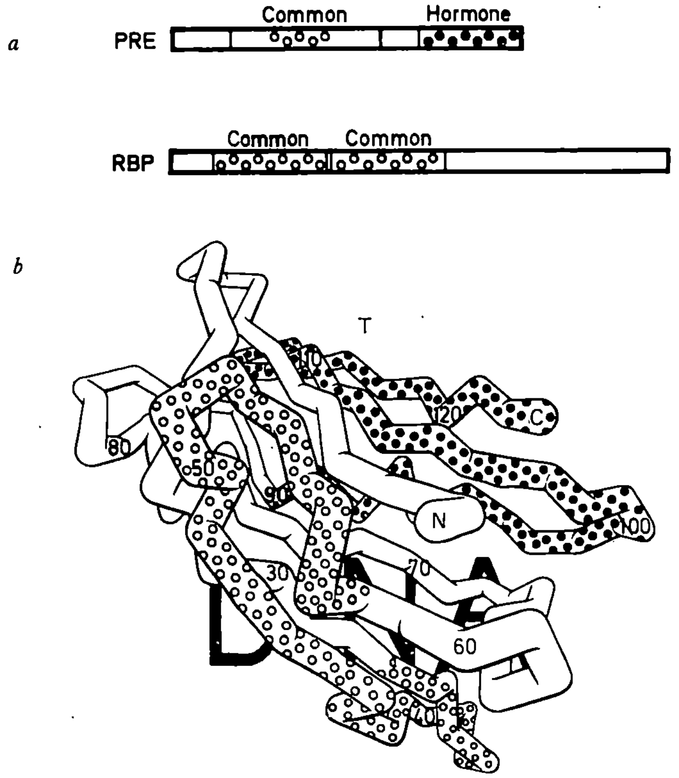


Fig. 3 *a*, Schematic representation of the regions discussed in prealbumin (PRE) and retinol-binding protein (RBP). 'Hormone' is the region homologous to the glucagon–secretin group of peptides; the other three homologous regions are indicated by 'Common'. *b*, Illustration of the regions discussed in a drawing of the conformation of the prealbumin subunit. Filled dots indicate the 'Hormone' region and empty dots the central part of the 'Common' region (positions in Fig. 2). In the native tetramer, a second subunit is above the one shown, with the binding site for thyroxine (indicated by T) in between, and another subunit is behind the one shown with the DNA binding site (indicated by DNA) in between. This drawing was made with the help of Fig. 3*b* of ref. 8.

The sequence similarities observed for prealbumin suggest evolutionary connections with gastrointestinal prohormones and retinol-binding protein. The data are compatible with the known properties of the prohormones, the conformation and binding interactions of prealbumin, and with previous ideas about the origins of receptors. They also suggest another superfamily of related proteins, indicating complex endocrine functions while helping to explain the unusual properties of prealbumin.

We thank Hans Nilsson for the drawing, Dr Steven Feinmark for linguistic help, and the Swedish MRC and Cancer Society for support (projects 13X-3532, 13X-1010 and 620, respectively). *Note added in proof*: Another comparison involving retinol-binding protein has just been reported[27] in which an internal although different duplication was observed (only second in present comparison). The discrepancies illustrate difficulties in comparisons and may question conclusions, but could also indicate further overlapping duplications.

Received 1 December 1980; accepted 25 February 1981.

1. Gitlin, D. & Gitlin, G. D. in *The Plasma Proteins* 2nd edn (ed. Putnam, F. W.) 264–317 (Academic, New York, 1975).
2. Gonzalez, G. thesis, Univ. Oxford (1972).
3. Gonzalez, G. & Offord, R. E. in *Atlas of Protein Sequence and Structure* (ed. Dayhoff, M. O.) Vol. 5, Suppl. 1, 81 (Silver Spring, Maryland 1973).
4. Kanda, Y., Goodman, D. A., Canfield, R. E. & Morgan, F. J. *J. biol. Chem.* **249**, 6796–6805 (1974).
5. Blake, C. C. F., Geisow, M. J., Swan, I. D. A., Rerat, C. & Rerat, B. *J. molec. Biol.* **88**, 1–12 (1974).
6. Blake, C. C. F. & Oatley, S. J. *Nature* **268**, 115–120 (1977).
7. Blake, C. C. F., Burridge, J. M. & Oatley, S. J. *Biochem. Soc. Trans.* **6**, 1114–1118 (1978).
8. Blake, C. C. F., Geisow, M. J., Oatley, S. J., Rérat, B. & Rérat, C. *J. molec. Biol.* **121**, 339–356 (1978).
9. Kanai, M., Raz, A. & Goodman, D. S. *J. clin. Invest.* **47**, 2025–2044 (1968),
10. Robbins, J. & Rall, J. E. *Physiol. Rev.* **40**, 415–489 (1960).
11. Pettersson, T. & Carlström, A. *Abstr. Eur. Congr. clin. Chem.* **3**, 16 (1979).
12. Pettersson, T. & Carlström, A. *Scand. J. clin. Lab. Invest.* (submitted).
13. Burton, P., Iden, S., Mitchell, K. & White, A. *Proc. natn. Acad. Sci. U.S.A.* **75**, 823–827 (1978).
14. Jacobsson, B., Pettersson, T., Sandstedt, B. & Carlström, A. *IRCS Med. Sci.* **7**, 590 (1979).
15. Dayhoff, M. O. (ed.) *Atlas of Protein Sequence and Structure* Vol. 5, Suppl. 2, 125–126 (Silver Spring, Maryland, 1976).
16. Jörnvall, H. *et al. FEBS Lett.* **123**, 205–210 (1981).
17. Tager, H. S. & Steiner, D. F. *Proc. natn. Acad. Sci. U.S.A.* **70**, 2321–2325 (1973).
18. Moody, A. J., Jacobsen, H. & Sundby, F. in *Gut Hormones* (ed. Bloom, S. R.) 369–378 (Churchill Livingstone, Edinburgh, 1978).
19. Rask, L., Anundi, H. & Peterson, P. A. *FEBS Lett.* **104**, 55–58 (1979).
20. Goodman, R. H. *et al. Proc natn. Acad. Sci. U.S.A.* **77**, 5869–5873 (1980).
21. Brew, K., Vanaman, T. C. & Hill, R. L. *J. biol. Chem.* **242**, 3747–3749 (1967).
22. Kurosky, A., Barnett, D. R., Rasco, M. A., Lee, T. -H. & Bowman, B. H. *Biochem. Genet.* **11**, 279–293 (1974).
23. Hunt, L. T. & Dayhoff, M. O. *Biochem. biophys. Res. Commun.* **95**, 864–871 (1980).
24. Adelson, J. W. *Nature* **229**, 321–325 (1971).
25. Seligman, P. A. & Allen, R. H. *J. biol. Chem.* **253**, 1766–1772 (1978).
26. Kouvonen, I. & Gräsbeck, R. *Biochem. biophys. Res. Commun.* **86**, 358–364 (1979).
27. Rask, L. *et al. Ann. N.Y. Acad. Sci.* (in the press).

Structure of the binuclear iron complex in metazidohaemerythrin from *Themiste dyscritum* at 2.2. Å resolution

R. E. Stenkamp, L. C. Siecker & L. H. Jensen

Departments of Biological Structure and Biochemistry, University of Washington, Seattle, Washington 98195, USA

J. Sanders-Loehr

Department of Chemistry, Portland State University, Portland, Oregon 97207, USA

Haemerythrin and myohaemerythrin are non-haem iron proteins utilized by several marine invertebrates in transporting and storing oxygen. Haemerythrin is oligomeric [usually octameric, but also trimeric[1,2], and tetrameric and dimeric (C. A. Appleby, personal communication)] whereas myohaemerythrin is a monomeric analogue of the haemerythrin subunit. Physicochemical investigations[3-5] of these molecules have shown that (1) each subunit binds two iron atoms by means of amino acid side chains, (2) one molecule of O_2 is bound to each binuclear iron centre, and (3) the iron atoms are in the +3 oxidation state in oxy and met forms of the proteins and +2 in deoxy. Models have been proposed for the complexes in haemerythrin and myohaemerythrin on the basis of crystallographic studies of the proteins. A μ-oxo bridged model (Fig. 1*a*) has been proposed from the results of a crystallographic analysis of metazidomyohaemerythrin from *Themiste zostericola*[6]. Based on a similar analysis of methydroxohaemerythrin from *Themiste dyscritum*, previously referred to as metaquohaemerythrin[7], we proposed a confacial bi-octahedron model with no atom identified as a μ-oxo bridge (Fig. 1*b*)[8]. Previously[9], we stated that "at least one of the models is incorrect". In fact, neither completely accounts for the electron density from the present 2.2 Å resolution difference map of the complex and surrounding electron density of metazidohaemerythrin from *T. dyscritum*.

Crystals of metazidohaemerythrin were obtained by direct reaction of oxyhaemerythrin with 0.01 M NaN_3 and subsequent crystallization by dialysis against low ionic strength buffer. The space group is P4, $a = b = 86.61(2)$ Å, $c = 80.90(2)$Å. The 2.2-Å data set, ~60,000 reflections, including Friedel pairs, was collected by the $\omega/2\theta$ step-scan method[10] from three metazidohaemerythrin crystals (27,131 reflections with $I > 2\sigma(I)$, 31,532 total reflections). Both empirical absorption[11] and deterioration corrections were applied.

A previously refined model of methydroxohaemerythrin[12] served as the initial model for this study. All atoms associated with the iron complex were removed, including those beyond the β-carbon atoms in the side chains of the amino acids shown as ligands in Fig. 1*b*. For the model lacking the complex, R $(= \Sigma||F_{o,azide}| - |F_c||/\Sigma|F_{o,azide}|)$ was 0.288. This model was sub-

0028-0836/81/210263—02$01.00

jected to two cycles of restrained least-squares refinement[13] to relieve the calculated phases of bias from the input structure, reducing R to 0.246. A difference electron density map calculated with coefficients $|F_{o,\text{azide}}|-|F_c|$ and calculated phases provided an essentially unbiased view of the electron density associated with the complexes in each of the four crystallographically independent subunits. Features of the difference density, such as resolution of the iron atoms, resolution across the ring of tyrosine 109 and resolution across the six-membered rings formed by the carboxylate groups, the iron atoms and the bridging atom, indicate the high quality of the map.

A straightforward interpretation of the difference density map for metazidohaemerythrin leads to the model shown in Fig. 2 having features in common with each of the earlier ones. The complex consists of two octahedrally coordinated iron atoms bridged by two carboxylate groups and an oxygen atom. The other ligands are three histidine side chains for one iron and two histidine side chains and the azide ion for the other. There is no density connecting tyrosine 109 (114 in myohaemerythrin) and the iron complex. The azide ion clearly binds to only one iron atom, rather than bridging the two as suggested by our earlier, low-resolution metazido study[14]. The observed mode of binding is as predicted by Gay and Solomon in interpreting their polarized, single-crystal spectroscopic results[15]. The coordination portrayed in Fig. 2 has not been observed in any other iron complex or metalloprotein, although it is similar to several small molecule iron carboxylate structures[16].

The model proposed for the complex in metazidomyohaemerythrin and shown in Fig. 1*a* has the bridging oxygen atom (a feature supported by spectroscopic and magnetic studies), but lacks the bridging carboxylate groups. On the other hand, the model proposed for the complex in methydroxohaemerythrin shown in Fig. 1*b* has bridging aspartic and glutamic acid side chains, the details of which were left unspecified. The model also has an additional bridging group which was assumed to be a water molecule or a hydroxide ion. Crystallographic refinement of the methydroxohaemerythrin form at 2.0 Å resolution had previously led to a more detailed view of the complex showing that the carboxylate groups bridge the two iron atoms with separate oxygen atoms on each iron. Tyrosine 109 was found to be too far from the nearest iron atom to be a ligand although it remained in the same eclipsed position observed earlier[8]. It is not coordinated either directly or indirectly through a water molecule, implying that one of the iron atoms is penta-coordinate. The resulting model is essentially the same as the present metazido model shown in Fig. 2, but lacks a ligand at the azide binding site.

A major question posed by the metazido structure and the bridging oxygen is whether the bridging atom in methydroxo is a single oxygen atom or some other species such as hydroxide. Another is whether there is any evidence for a hydroxide ion being bound at the azide binding site. In answer to the first question, we note that it is impossible in the electron density map at the present resolution to differentiate between a single oxygen atom, a hydroxide ion or a water molecule. The only ligand density not accounted for by amino acid side chains in methydroxohaemerythrin was at the bridging position. In accordance with the generally held opinion at the time that in met derivatives, small molecule ligands such as OH^-, N_3^-, SCN^- and H_2O were bound directly to the binuclear iron complex[17–19], we considered the bridging density to be the small ligand. However, as the bridging group in the metazido structure is present in the same location, it seems to be a common feature of the complex and is probably a μ-oxo atom as indicated by spectroscopic and magnetic evidence.

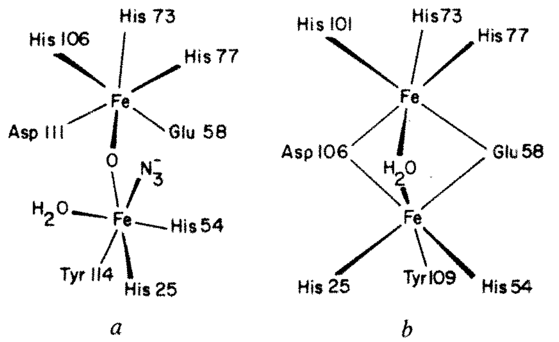


Fig. 1 Earlier models proposed for the binuclear iron complex in (*a*) metazidomyohaemerythrin from *T. zostericola*[6] and (*b*) methydroxohaemerythrin from *T. dyscritum*[8] (drawn from Fig. 4, ref. 9).

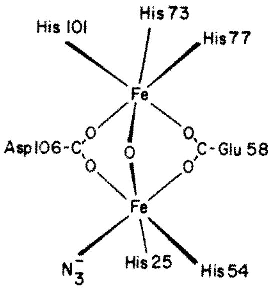


Fig. 2 The iron complex in metazidohaemerythrin from *T. dyscritum*.

To answer the second question and to make a more definitive comparison of the methydroxo and metazido forms of haemerythrin, we have restricted the resolution of the methydroxo data to 2.2 Å and treated it in the same way as the metazido data. We find small peaks in the difference density at the azide binding sites in all four subunits in the asymmetric unit although they are no larger than some noise peaks found in the open solvent regions of the crystal. The low occupancy of the sites indicates that only a fraction of the molecules have a sixth (as yet unidentified) ligand. Thus, the major form of the complex in the conditions pertaining in the crystals has an unexpected penta-coordinate iron atom.

The present model clarifies the contradictory crystallographic models previously proposed for the complexes in haemerythrin and myohaemerythrin. The difficulty arose in part because the met derivatives had been assumed to be similar, but significant structural differences clearly exist between at least two of them. Differences are observed in the spectra of the various states of these molecules, and high-resolution studies of other derivatives of met, oxy and deoxy forms of haemerythrin may allow significant correlations between the spectroscopic and structural results.

We thank W. A. Hendrickson, J. Smith, T. M. Loehr and I. M. Klotz for the exchange of data, information and viewpoints concerning the structure of these molecules. This work has been supported by NIH grant GM-10828.

Received 19 January; accepted 24 February 1981.

1. Liberatore, F. A., Truby, M. F. & Klippenstein, G. L. *Archs Biochem. Biophys.* **160,** 223–229 (1974).
2. Addison, A. W. & Bruce, R. E. *Archs Biochem. Biophys.* **183,** 328–332 (1977).
3. Okamura, M. Y. & Klotz, I. M. *Inorganic Biochemistry* (ed. Eichorn, G. L.) 320–343 (Elsevier, New York, 1973).
4. Klotz, I. M., Klippenstein, G. L. & Hendrickson, W. A. *Science* **192,** 335–344 (1976).
5. Loehr, J. S. & Loehr, T. M. *Adv. inorg. Biochem.* **1,** 235–252 (1979).
6. Hendrickson, W. A. *Naval Res. Rev.* **31,** 1–20 (1978).
7. Dunn, J. B. R., Addison, A. W., Bruce, R. E., Loehr, J. S. & Loehr, T. M. *Biochemistry* **16,** 1743–1749 (1977).
8. Stenkamp, R. E., Sieker, L. C. & Jensen, L. H. *Proc. natn. Acad. Sci. U.S.A* **73,** 349–351 (1976).
9. Stenkamp, R. E. & Jensen, L. H. *Adv. inorg. Biochem.* **1,** 219–233 (1979).
10. Hanson, J. C., Watenpaugh, K. D., Sieker, L. & Jensen, L. H. *Acta crystallogr.* **A35,** 616–621 (1979).
11. North, A. C. T., Phillips, D. C. & Mathews, F. S. *Acta Crystallogr.* **A24,** 351–359 (1968).
12. Stenkamp, R. E., Sieker, L. C. & Jensen, L. H. *American Crystallographic Association*, Summer Meet., Calgary, Abstr. (1980).
13. Konnert, J. H. *Acta crystallogr.* **A32,** 614–617 (1976).
14. Stenkamp, R. E., Sieker, L C. & Jensen, L. H. *J. molec. Biol.* **126,** 457–466 (1978).
15. Gay, R. R. & Solomon, E. I. *J. Am. chem. Soc.* **100,** 1972–1973 (1978).
16. Thundathil, R. V., Holt, E. M., Holt, S. L. & Watson, K. J. *J. Am. chem. Soc.* **99,** 1818–1823 (1977).
17. Keresztes-Nasy, S. & Klotz, I. M. *Biochemistry* **4,** 919–931 (1965).
18. Darnall, D. W., Garbett, K. & Klotz, I. M. *Biochem. biophys. Res. Commun.* **32,** 264–271.
19. Garbett, K., Darnall, D. W. & Klotz, I. M. *Archs Biochem. Biophys.* **142,** 455–470 (1971).

MATTERS ARISING

Si, Al ordering in Linde A zeolite

ALTHOUGH we congratulate Bursill *et al.*[1] on their excellent high-resolution images of zeolite Linde A, we wish to question some of their interpretation—in particular the evidence on Si, Al ordering.

For a Linde A zeolite with Si/Al ratio of exactly 1, regular alternation of Si and Al leads inexorably to a super-cell with $a = 2.4$ nm and space group $Fm\bar{3}c$ from the pseudo-cell with $a = 1.2$ nm and space group $Pm\bar{3}m$ for fully disordered Si and Al. Although many structure determinations have been made with the pseudo-cell, the superstructure diffractions are easily observable on long-exposure X-ray photographs. Gramlich and Meier[2] obtained a satisfactory refinement of the superstructure of Zeolite A in which alternate tetrahedra had dimensions appropriate for occupancy at least mainly by Si and by Al. Recent refinements of dehydrated K-exchanged zeolite A (ref. 3) and dehydrated zeolite A (ref. 4) (the as-synthesized Na variety) were made to test the concept of zero-coordination, and satisfactory refinements were obtained of the superstructure. All these X-ray diffraction data, and much unpublished data, cannot be explained by the ordering model proposed by Engelhardt *et al.*[5] on the basis of high-resolution magic-angle spinning ^{29}Si NMR measurement. This model yields a unit cell with $a = 1.2$ nm and space group Pm3, and can be ruled our immediately by the X-ray diffraction data which require a cell with $a = 2.4$ nm. In the following, note that zeolite A has strong alternation of Si and Al on the tetrahedral nodes of the framework, and that only a small percentage of nodes are not surrounded by four atoms of the opposite type.

Pluth and Smith[3] found evidence from detailed electron microprobe analyses of A crystals made by Charnell's method that the Si/Al ratio is >1 (~ 1.03). This was supported by population estimates from published[3,4] and unpublished X-ray measurements which indicated less than 12 exchangeable monovalent cations in the pseudo-cell (for example 11.5 Na atoms in dehydrated zeolite A (ref. 4)). Such deviation from the ideal value must cause some Si atoms to be surrounded by less than four Al atoms. Silica-rich regions might have a different stability than regions with Si/Al equal to unity, and Bursill *et al.* might find it useful to measure the Si/Al ratio of their crystals.

Although Pluth and Smith refined their structures with the 2.4-nm cell and space group $Fm\bar{3}c$, they and earlier workers found several observed diffractions which violated the c glide plane. Further measurements are now in progress to discover the source of these violations. Note that these diffractions are weak and might result from the exchangeable cations and not from the Si and Al atoms. In the dehydrated K-exchanged- and Na-A varieties, some exchangeable cations have a low occupancy of sites on the walls of cavities, and local distortions of the framework can be expected. Such distortions must be considered as a potential cause of space group violations.

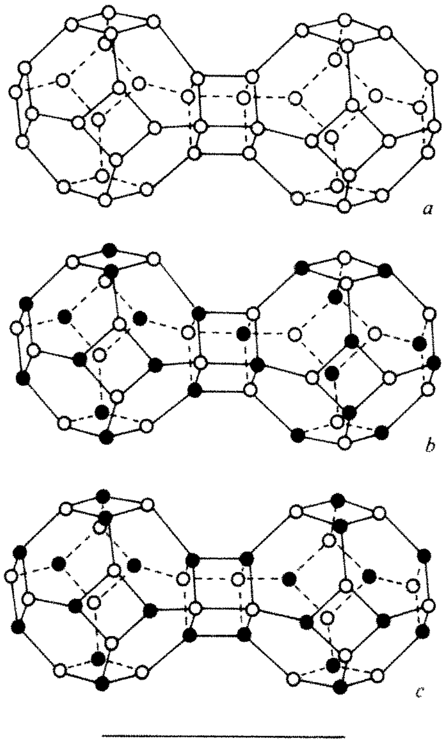


Fig. 1 Three ordering schemes for zeolite A. *a*, Complete disorder of tetrahedrally coordinated cations would give a 1.2-nm cell repeat with space group $Pm\bar{3}m$. *b*, Regular alternation of Si and Al requires 2.4 nm and $Fm\bar{3}c$. *c*, The isometric cell for the model of Engelhardt *et al.*[5] has a 1.2-nm repeat and Pm3; note the absence of a mirror plane in ⟨110⟩ directions.

The ^{29}Si NMR results were interpreted by Engelhardt *et al.*[5] in terms of an ordering scheme in which each Si is surrounded by three Al and one Si, and each Al by three Si and one Al (Fig. 1). This ordering scheme gives an isometric cell with $a = 1.2$ nm and space group Pm3, in contrast with 1.2 nm ($Pm\bar{3}m$) for complete disorder and 2.4 nm ($Fm\bar{3}c$) for regular alternation of Si and Al. The 3:1 ordering scheme is not compatible with the 2.4-nm cell repeat found consistently in long-exposure X-ray photographs, and can be ruled out except for small regions. Simple comparison of NMR data from one zeolite with another is obviously unreliable, and further interpretation is needed of the data in ref. 5. The observation by Bursill *et al.* of hkl (all odd) reflections with $h + l = 2n$ does not confirm "the recent criticism of space group Fm3c". This class of diffractions obeys the F centred lattice deduced by the X-ray crystallographers. Violations of the condition for the c glide plane (l must be even in hhl and $0kl$) are needed to cast doubt on $Fm\bar{3}c$ (for example, observation of 111; 3,3,17; 7,7,23; 13,13,11; 15,15,15 for dehydrated A ref. 4).

We suggest that Linde A zeolite has strong alternation of Si and Al atoms, and that only local regions deviate therefrom. The new refinement of dehydrated A zeolite[4] could be used to make a computer-simulated image for comparison with Fig. 1 of ref. 1: qualitatively a good match is likely, thereby indicating extraction of water by the pre-pumping procedure[1]. We look forward to more important observations of zeolites by electron microscopy, and suggest that existing X-ray diffraction data can provide a fundamental basis for interpretation.

We acknowledge grant NSF CHE 78-26579, and thank E. Flanigen of the Linde Division of Union Carbide Corporation for review.

JOSEPH V. SMITH
JOSEPH J. PLUTH

Department of the Geophysical Sciences,
Universtiy of Chicago,
Chicago, Illinois 60637, USA

1. Bursill, L. A., Lodge, E. A. & Thomas, J. M. *Nature* **286,** 111–113 (1980).
2. Gramlich, V. & Meier, W. M. *Z. Kristallogr.* **133,** 134–149 (1971).
3. Pluth, J. J. & Smith, J. V. *J. phys. Chem.* **83,** 741–749 (1979).
4. Pluth, J. J. & Smith, J. V. *J. Am. chem. Soc.* **102,** 4704–4708 (1980).
5. Engelhardt, G., Zeigan, D., Lippmaa, E. & Mägi, M. *Z. anorg. allg. Chem.* **468,** 35–38 (1980).

BURSILL *ET AL.* REPLY—We thank Smith and Pluth for their comments but we are not convinced by their arguments. Recent experimental results reaffirm that Na-A zeolites possess an aluminosilicate ordering scheme, in which each Si ion is connected through oxygen bridges to three Al ions and one Si ion, and each Al ion to three Si and one Al; that is 3:1, rather than 4:0, ordering prevails.

Our reasons for disputing the accepted space group $Fm\bar{3}c$ rests partly on some of the X-ray results of Smith and Pluth[1,2] but chiefly on results from three other, independent techniques: electron diffraction[3–5], high-resolution magic-angle spinning ^{29}Si NMR[6–8], and neutron diffraction[9]. Weak superlattice reflections appear readily in electron diffraction patterns, requiring exposure times of ~ 1 s. enabling many hundreds of crystals

0028-0836/81/210265—02$01.00

to be studied. We established that our preparations are homogeneous, with no evidence for local variations in intensity of weak reflections, and that doubling is observed in all zones when it is expected for a 24.6-Å unit cell. In particular *hhl* (inadvertently printed as *hkl* in ref. 3) reflections satisfying $h + l = 2n$ were observed. As it is not easy to assign unambiguously space groups using electron diffraction alone[4] we have sought further information concerning space group assignment. Surprisingly, rather cogent evidence that the space group is not $Fm\bar{3}c$ comes from Smith and Pluth who meaasured 111; 3,3,17; 7,7,23; 13,13,11 and 15,15,15 intensities for dehydrated Na-A (ref. 1) and 999; 31,1,1, 21,21, 9 for dehydrated K-A (ref. 2). All of these reflections contradict the c-glide operation of $Fm\bar{3}c$. Smith and Pluth suggest that these reflections arise from displacements of cations and/or framework atoms, dismissing the NMR results[6–8], which strongly implicate 3:1 ordering, on the grounds "that simple comparison of NMR data from one zeolite sample with another is obviously unreliable" and that the model proposed[6,7] does not lead to a 24.6-Å cell. The NMR results are especially significant as careful calibrations of the chemical shifts, using known aluminosilicates as standards[8], show the technique is extremely sensitive to the Si–O–Si (Al) coordination. The technique yields results for bulk specimens and merits at least the respect accorded to a crystal structure determination on one or two carefully selected X-ray specimens. We have pointed out[4] that the model proposed by Engelhardt *et al.*[6,7], based on 3:1 ordering, yields an isometric cell with $a = 12.3$ Å and space group Pm3, and suggested that further NMR and diffraction studies, encompassing a range of Si/Al ratios, should prove instructive, as cell doubling may arise to accommodate excess Al (or Si)[4,5].

New data on Si, Al ordering in zeolites have emerged from neutron diffraction studies, carried out in collaboration with A. K. Cheetham. Three dehydrated samples, covering a range of Si/Al ratios, were studied. Our analysis shows that the neutron diffraction patterns cannot be indexed using space group $Fm\bar{3}c$. However, refinement of the structure is proceeding saisfactorily using a space group having rhombohedral symmetry. Analysis of the geometrically possible Si, Al distributions required by this space group, and which maintain 3:1 ordering, required by the NMR results, and which give 24.6-Å cells, as required by the X-ray and electron diffraction results, leads to a new description of the structure, details of which are to be given elsewhere[9].

We believe, therefore, that the 4:0 ordering scheme advocated by Smith and Pluth for Linde A must be viewed with reservation, at least for the bulk of Na-A preparations. Application of the above-mentioned range of techniques to other zeolites, where novel ordering schemes may be uncovered, is indicated.

L. A. BURSILL*
ELIZ. A. LODGE
J. M. THOMAS

Department of Physical Chemistry,
University of Cambridge,
Lensfield Road,
Cambridge CB2 1EP, UK

[*] Permanent address: School of Physics, University of Melbourne, Parkville, 3052, Australia.

1. Pluth, J. J. & Smith, J. V. *J. Am. chem. Soc.* **102,** 4704 (1980).
2. Pluth, J. J. & Smith, J. V. *J. phys. Chem.* **83,** 741 (1979).
3. Bursill, L. A., Lodge, E. A. & Thomas, J. M. *Nature* **286,** 113 (1980).
4. Lodge, E. A., Bursill, L. A. & Thomas, J. M. *JCS Chem. Commun.* 856 (1980).
5. Bursill, L. A., Lodge, E. A. & Thomas, J. M. (submitted).
6. Engelhardt, G., Zeigan, D., Lippmaa, E. & Magi, M. *Z. anorg. allg. Chem.* (in the press).
7. Engelhardt, G., Kanath, D., Samonson, A. & Tarmak, M. *Adsorption of Hydrocarbons in Zeolites* (Berlin Aldershoft, 1979).
8. Lippmaa, E., Magi, M., Samonson, A., Engelhardt, G. & Grimmer, A.-R. *J. Am. chem. Soc.* **102,** 4889 (1980).
9. Bursill, L. A., Lodge, E. A., Thomas, J. M. & Cheetham, A. K. (submitted).

Modern adaptations in orang-utans?

SMITH AND PILBEAM'S[1] letter, proposing that orang-utan sex differences—body size and certain features of dental and craniofacial morphology—cannot be explained by the species' modern adaptations and thus are 'remnants' of a more terrestrial Pliocene pattern, states that sexual selection among orang-utans is not clearly indicated by field observations, citing Horr[2] to the effect that "direct evidence for male competition in the form of dominance or aggressive encounters is limited".

The Tanjung Puting study[3], now in its ninth year, indicates that all adult male-male encounters involve either aggression or avoidance as do most contacts between sub-adult and adult males. Aggression is the invariable response when two adult males encounter one another in the presence of an oestrous female. The large size of orang-utan males seems to be the result of sexual selection as there is intense male–male competition for oestrous females and female selection of males during consortships[3,4]. In addition, several workers have found orang-utan males and females using differential proportions of resources[3–5] and at least one[3] has argued for an ecological separation of the sexes which may help explain sex differences as well.

Several features of orang-utan dental and craniofacial morphology are at variance with those characteristic of other arboreal frugivores—this can be explained by the fact that orang-utans spend an average of 40% of their feeding time on bark, young leaves, insects and other foods[3]. During some months orang-utans spend less than 20% of their foraging time consuming fruit. One instance of meat-eating among a wild orang-utan population has just been reported (Sugardjito, personal communication). Furthermore, some of the fruits (for example *Mezzettia leptopoda*) eaten by orang-utans are extremely hard. Chimpanzees use tools to open very hard fruits[6], but orang-utans spend hours opening hard nuts with their molars, which may account for their exceptionally thick molar enamel.

Finally, note that orang-utans can be quite terrestrial, bringing them closer to chimpanzees in this regard than is usually realized[3]. At Tanjung Puting adult males averaged 66 min per day on the ground with one male spending up to 6 h per day on the forest floor. A clear understanding of modern orang-utan adaptation suffices to explain most morphological features cited by Smith and Pilbeam as indicative of a more terrestrial Pliocene pattern. In regards to the fossil evidence for terrestrial patterns, we need post-cranial remains of ancestral orang-utans before we can discuss such patterns.

B. M. F. GALDIKAS

Orang-utan Research and
Conservation Project,
Pangkalan Bun,
Kalimantan Tengah,
Indonesia

1. Smith, R. J. & Pilbeam, D. R. *Nature* **284,** 447–448 (1980).
2. Horr, D. A. *Primate Behavior* Vol. 4 (ed. Rosenblum, L.A.) 307–323 (Academic, New York, 1975).
3. Galdikas, B. M. F. thesis, Univ. California, Los Angeles (1978). Galdikas, B. M. F. *The Great Apes* (eds Hamburg, D. A. & McCown, E. R.) 194–233 (Benjamin/Cummings, Menlo Park, 1979).
4. Rijksen, H. D. *A Fieldstudy on Sumatran Orang-utans* (Veenman & Zonen, Wageningen, 1978).
5. Rodman, P. S. *Feeding and Ranging Behaviour of Lemurs, Monkeys and Apes* (ed. Clutton-Brock, T. H.) 383–413 (Academic, London, 1977).
6. Teleki, G. *J. hum. Evol.* **3**, 575–594 (1974).

SMITH AND PILBEAM REPLY—The data reported by Galdikas are interesting, but the implications for our argument are unclear. What is directly known of ancestral orang-utans (their teeth and something of their distribution) is suggestive of a more terrestrial niche than that usually described for the extant species. Whether or not features of the extant animal support this hypothesis is secondary because one would expect the modern species to be more or less adapted to what it does. Also, whether or not the behaviours observed by Galdikas are sufficient to explain the morphological features in question is entirely speculative.

R. J. SMITH

Department of Orthodontics,
Dental School,
University of Maryland,
Baltimore, Maryland 21201, USA

D. R. PILBEAM

Department of Anthopology,
Yale University,
New Haven,
Connecticut 06520, USA

Biology and social science: wedding or rape?

Edmund Leach

Genes, Mind, and Culture: The Coevolutionary Process. By Charles J. Lumsden and Edward O. Wilson. Pp.512. ISBN 0-674-34475-8. (Harvard University Press: 1981.) $20. The book will be published in the UK in July, price £12.

THIS book comes so close to being a parody of the genre to which it belongs that I have had difficulty in believing that it is not intended as an academic hoax. Edward Wilson has made important contributions to the study of the social behaviour of insects and has achieved a wider notoriety as popularizer of the concept of sociobiology. He is also the author of a notably undistinguished book entitled *On Human Nature* (Harvard University Press, 1978); Lumsden is "a physicist who has extended his research into theoretical biology". The authors jointly claim an "abiding respect for social theory, the neurosciences and psychology". I cannot comment on their knowledge of the neurosciences and psychology but, on the evidence here presented, they appear to have no understanding of social theory whatsoever.

The basic proposition of the book stems from a belief that "the advance of men over chimpanzees and other higher primates was attained during an exceptionally rapid burst of evolution . . . over a period of approximately three million years". This alleged rapidity of a postulated unilinear evolutionary progress is held to be causally associated with a special kind of feedback reinforcement between the genetic constitution of human beings and their culture (here renamed "euculture"). This latter theme is expounded with the aid of a massive battery of diagrams, graphs and sophisticated-looking algebraic equations.

From my own doubtless prejudiced but none the less professional point of view, the whole of this elaborate apparatus is devoid of meaning because it assumes that human culture can be broken down into clusters of traits (which the authors call "culturgens") which can then be counted and subjected to statistical analysis.

The notion of culturgens is defined at some length and is summarized as "equivalent to the artifact type employed in archaeology (by D. L. Clarke)". From then on the whole argument presupposes that culturgens are readily identifiable entities. With very few exceptions, contemporary social and cultural anthropologists are unanimous in holding that human culture cannot be broken down into units of this or any other kind. The late D. L. Clarke, who was once my pupil, would certainly have rejected the idea.

However, empirical evidence plays only a secondary role in the authors' presentation. It is only drawn upon to illustrate points which they already judge to be self-evident. A typical example of the way this is done comes on p.6:

> From the time of the beginnings of a materials-based culture in the ancestral *Australopithecus* to the Upper Palaeolithic era of modern *Homo sapiens* the brain tripled in size. The cranial capacity of *Australopithecus* was 400–500 cc., comparable to that of the chimpanzee and gorilla. Two million years later *Homo erectus* had attained a capacity of 1000cc. Another million years saw an increase to 1400–1700 cc. in Neanderthal man and 900–2000 cc. in modern *Homo sapiens.*

Yet the authors must certainly know: (i) that is is far from clear that any of the Australopithecine species so far identified was directly ancestral to *Homo;* (ii) that *Australopithecus,* though man-like in shape, was very small, so it is not surprising that it had a small head; (iii) that the evidence available concerning the early hominids and the early varieties of *Homo* is statistically miniscule (in particular, as the authors' figures indicate, it is quite possible that all the *Homo* skulls fall within the range found in modern man); (iv) that since, in modern man, there is no correlation at all between cranial capacity and manifested intellectual capacity the size of our heads has no bearing whatever on the theme of the book which is "evolutionary advance" (whatever that means).

The "parody of science" aspect of the presentation can be illustrated from Chapter II at the end of which we are told that:

> Epigenesis is the total process of interaction between genes and environment during the course of development with the genes being expressed through epigenetic rules.

This explains the otherwise baffling opening sentences which read:

> The key element in the theory of gene–culture coevolution is the role of the epigenetic rules in culturgen choice. These rules and the bias curves they generate can serve as the molecular units in postulational-deductive models of the social sciences.

If we dig through the jargon it turns out that all this really means is that man's genetic endowments make it likely that some aspects of culture will be more elaborately developed than others. Thus visual and auditory discriminations, as manifested in speech and music and lexicon, are always more precise than those which depend upon taste and smell.

These platitudes about sensory perception are then extended into the field of custom and the concept of culturgens comes into its own. Once again the crass idiocy of the whole proceedings can only be illustrated by quotation:

> Female puberty initiation rites originate most frequently in societies where a young girl continues to reside at least half the time in her mother's household (in other words where uxorilocal or bilocal residence rules are in force), as well as in societies that depend heavily on the work of girls and young women to make a major contribution to sustenance activities.
>
> Recognising that some culturgens are more derivative and communicative in function than others, it is possible to delineate what appear to be chains of causation connecting the principal habitat, economic strategy, and both primary and derivative cultural choices. A very reduced example is the following:

Habitat and economy	Principal culturgens	Derivative culturgens
Desert herders	→ Polygyny, patrilineality, patrilocality	→ Absence of formal female initiation ceremony

Just what this could have to do with genetics is hard to imagine, unless of course the authors suppose that uxorilocal and patrilocal populations are racially distinct! But, for the record, it deserves note that the first half of the first paragraph is gibberish while the second half is such a let out that it would cover almost anything. "Female puberty initiation rites" is a broad category but the most drastic of the better known forms are those which involve clitoridectomy. These practices were traditionally confined to regions of East Africa where nomadic pastoralists with a more or less patrilineal/patrilocal/polygynous form of social organization are particularly prominent. It would be difficult to imagine a worse exemplification of the Lumsden–Wilson "chains of causation".

All the other examples of "gene–culture translation" which the authors mention (and they mention very few) are similarly defective. For example they claim that there is:

> a nearly universal avoidance of marriage and full sexual relations between full brothers and sisters. The epigenetic rule appears well

established: a deep sexual inhibition develops between those who live in close domestic contact during the first six years of life.

But no such rule has been "well established". On the contrary there is good evidence that in modern Western society sibling incest is quite common and there is no good reason to suppose that it is rare elsewhere. Moreover, in direct contradiction to what is asserted by Lumsden and Wilson, there is unambiguous evidence that monogamous marriage was widely practised in ordinary peasant families in Egypt during the first three centuries of our era and very probably for millennia before that. It does not require "epigenetic rules" to explain why marriages of this sort are usually forbidden; the proscription follows directly from the social theory for which Lumsden and Wilson have such abiding respect but so little understanding.

The human species is not genetically uniform but the variations seem to be minor. No correlation has yet been established between known genetic variations and morphological variations. Equally there is no clear correlation between morphological variations (the so-called "racial types"), cultural variations and/or modes of subsistence. But equally there are no cultural universals except at the level of such generalizations as "all human beings have speech", "all human societies distinguish between domestic space and non-domestic space".

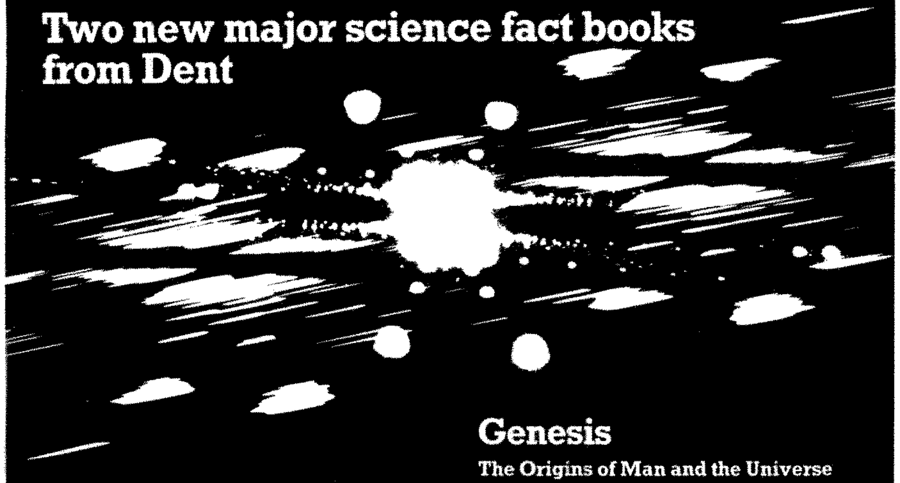
If epigenetic rules have had consequences for human evolution, it is difficult to see what sort of rules they might have been. On this central feature of the authors' argument the phoney mathematical apparatus which characterizes this text offers no illumination.

The book ends by setting a programme for social science:

> The decomposition of social behaviour into objective functional units, the discovery of new and sometimes surprising epigenetic rules, the measure of human genetic diversity, the tracing of the microevolution of the features of behavioral epigenesis, the paleobiological reconstruction of the origins of culture, the retrodiction [*sic*] of ethnography and connection of the results to new and more powerful covering laws in economics and sociology, perhaps even a sighting down the world-tube of possible future histories–these are activities that will come increasingly to occupy the social sciences as the links between biology and the study of culture are more powerfully forged.

I will readily agree that the findings of biology and of social science need to be compatible if either are to rate as science. But what is called for is mutual respect on both sides. The jargon loaded take-over bid proposed by Lumsden and Wilson is just bunk. We need a marriage — not a rape. ☐

Edmund Leach is a Fellow of King's College, Cambridge.



Two new major science fact books from Dent

Genesis
The Origins of Man and the Universe
JOHN GRIBBIN
With the same fluency of exposition as he showed in his much-praised *Timewarps*, John Gribbin now brilliantly describes how the earth has evolved through space and time, concluding with a chapter about the destination of mankind in the light of our new-found knowledge.
'John Gribbin's best book. Very good indeed.' *Professor Martin Rees, Director of the Institute of Astronomy, University of Cambridge*
Illustrated with 8 pages of photographs and 15 diagrams £7.95

The Edge of Infinity
Beyond black holes to the end of the universe
PAUL DAVIES
In his latest book Paul Davies, author of the highly acclaimed *Other Worlds*, takes us beyond black holes to naked singularities and the destruction of spacetime. Naked singularities, unclothed by black holes, not subject to any known laws, represent the threshold to a strange new physical world.
Illustrated with 52 line drawings £7.95

Circle No. 19 on Reader Enquiry Card.

Marine mineralogy

Roger G. Burns

Underwater Minerals. By D. S. Cronan. Pp.362. ISBN 0-12-197480-4. (Academic: 1980.) £24.80, $57.50.

THE publication of J. L. Mero's *The Mineral Resources of the Sea* by Elsevier 16 years ago drew attention to the vast mineral deposits on the seafloor. Subsequent events — notably expanded oceanographic research, improved mineral beneficiation and extraction processes, real and threatened embargoes of industrial minerals and fuels, expansion of coastal state jurisdiction to the 200 mile limit, and legal ramifications of the United Nations Law of the Sea Conferences towards deep-sea mining of manganese nodules — have heightened the public's awareness of the potential of submarine mineral resources.

The appearance of Dr Cronan's book, therefore, is particularly significant at this time of rapidly accumulating information from modern oceanographic and marine geochemical research. In one concise volume, Cronan surveys several types of underwater mineral deposits, relates their origins to current global plate tectonic models, describes submarine exploration methods and explores some of the consequences of mining marine minerals. He calls upon his own experiences and observations from years of oceanographic cruises and laboratory research. The text focuses on various types of mineral deposits mainly located beneath the oceans; the extraction of minerals from seawater by modern processes or from ancient evaporite deposits is not discussed.

Chapters in Cronan's book are devoted to placers and aggregates, including deposits of gold, platinum, diamond, cassiterite (tin oxide), ilmenite and other heavy minerals mined near continental margins; phosphorites on continental shelves; miscellaneous authigenic minerals such as zeolites and barite; manganese nodules and encrustations from deep sea floors, marginal seas and lakes; metalliferous sediments related to volcanic activity along spreading centres, mid-ocean ridge crests and fracture zones, island arcs and submarine volcanoes; and sub-surface deposits such as ore and coal seams adjacent to continents, and sulphide mineralization in modern or uplifted (ophiolite) ocean floor basalts. The concluding three chapters relate marine mineral deposits to recent theories of plate

To help scientists cope with the complex problem of information retrieval, *The Use of Scientific Information Services* by R. Davidge and E.R. Wooding describes the use of information services in libraries, and contains a listing of reference sources. The report is available from The Library, Royal Holloway College, Egham Hill, Egham, Surrey TW20 0EX, UK, price £2.

tectonics and ocean evolution, describe methods for exploring underwater mineral deposits, and review some of the factors leading to the mining, and consequences resulting from exploitation, of seafloor minerals.

On the whole, the author does an excellent job in surveying such a broad subject. He references, perhaps uncritically, a wide spectrum of published papers including many of his own noteworthy contributions, so that a beginner is well informed about each topic while a researcher is guided to significant, albeit familiar, publications. Obviously, in one 362-page book, there is inevitable lack of depth of coverage in some topics. However, recent readily available specialist texts and monographs complement Cronan's book by providing detailed coverage of marine manganese nodule deposits and other specific marine minerals. The author has achieved his objective to write an up-to-date textbook of interest to undergraduate and postgraduate students, civil servants concerned with marine affairs and people generally interested in the marine environment. □

Roger G. Burns is Professor of Mineralogy and Geochemistry in the Department of Earth and Planetary Sciences at the Massachusetts Institute of Technology.

Stylish distillation for petrologists

M. P. Atherton

Metamorphic Petrology: Mineralogical, Field and Tectonic Aspects. 2nd Edition. By Francis J. Turner. Pp.524. ISBN 0-07-065501-4. (McGraw-Hill: 1981.) £25.95, $ 40.05.

THE second edition of Francis Turner's book is a considerable modification and updating of a volume, first published in 1968, which has been a milestone on the path towards our understanding of metamorphic rocks. It has increased in size from 403 to 524 pages, and although many of the original diagrams have been kept, many new ones have been added. Nevertheless, even with these additions the stylish character of the book has been maintained. There are eleven chapters as opposed to eight in the first edition, and most of this change has involved a new division and expansion of the subject, rather than the addition of entirely new material. The references are now placed at the end of the book rather than at the end of each chapter, and lie next to an extensive author index. It is interesting to note that even though many new authors are cited, Eskola, Fyfe, Verhoogen, Goldschmidt and Tilley still dominate the references.

In spite of the fact that much fresh material and many new ideas have been introduced, the main thread of the original book is maintained, and is "still by way of Eskola's facies concept" which allows the mineralogical and field material to be most readily "perceived, marshalled and analysed". The new edition also follows the approach in two earlier books (*Metamorphic Reactions and Metamorphic Facies* by W.S. Fyfe, F.J. Turner and J. Verhoogen *Geol. Soc. Am. Mem.* 73, 1958; *Igneous and Metamorphic Petrology* by F.J. Turner and J. Verhoogen, McGraw-Hill, 1960) both classics of their time, in that they were the first to present an integrated experimental and thermodynamic approach in the field of petrology. Thus it contains ACF, AKF, AFM diagrams indicating the chemographic relations of mineral assemblages defining the facies, PT diagrams, some simple thermodynamic theory and applications, and many maps of metamorphic areas showing the general geology and metamorphic zonal patterns.

In the account of these elements of metamorphism the book is excellent, and insofar as the aim of all metamorphic geologists as to be able to accurately plot the PT path of metamorphic rocks, the discussion on experimental data, Shreinemaker's diagrams and the use of thermodynamics is sensible and rightly critical — as is to be expected from such an experienced researcher and teacher whose working life has spanned the old and the new in metamorphic petrology. I particularly liked the addition of recent ideas on pore solutions, ionic reactions, Shreinemaker's analysis, geobarometry and geothermometry, burial metamorphism and new mineral equilibria.

However, if the aim of the metamorphic geologist is to plot the path of his rocks in PT space, I think it is also necessary to consider the textures in the rocks, to integrate them with the deformation history and with the sequence of reactions involving a fluid phase which occurred during metamorphism. This new edition includes little on these important aspects although, in fairness, much of the progress made in those areas, particularly in integrating them, was made after the first edition. Their incorporation into the present book would have involved a completely fresh approach, indeed a much longer book.

Despite these comments this new edition will, I am sure, continue to be influential and be much referred to. It should be read by every thinking student of metamorphism because it puts the subject into proper historical context and is the distillation of the lifetime views of a sharp-eyed, thoughtful and influential metamorphic petrologist. □

M. P. Atherton is Reader in Geology at the University of Liverpool.

Records of the past

John Currey

Skeletal Growth of Aquatic Organisms: Biological Records of Environmental Change. Topics in Geobiology, Vol.1. Edited by D.C.Rhoads and R.A. Lutz. Pp.750. ISBN 0-306-40259-9. (Plenum: 1980.) $47.50, £29.93.

THIS book explores what we can learn from the skeletal structure of aquatic organisms about the environment in which they lived, and how the environment affected their lives. The editors point out that until recently much of the drive for such studies has come from palaeontologists interested in reconstructing palaeoenvironments, but that now ecologists are also becoming interested in how the recent past can be revealed by skeletons.

The editors have clearly put much thought into the organization of the book, and have given their contributors clear briefs. As a result, they have produced a book that is admirable for the beginner and also for old hands.

There are 16 chapters and 12 appendices. Twelve of the main chapters concern molluscs; they range from a general overview of shell growth and form in bivalves to particular examples of the effect of the environment on shell structure of *Mercenaria*, and the recognition of developmental types from larval shell morphology. The balance in this mollusc section seems right; the major emphasis is on molluscs as a whole, and how the environment influences shell growth, yet there are some examples of how these principles apply to particular populations or species. In all the contributions it is reiterated that while the environment has marked effects on the shell, the interpretation of, for instance, check lines, changes in mineralogy, or oxygen isotope ratio, needs great care. The message, however, is not that the subject is so difficult that only people who have worked on it for 20 years should attempt to publish but rather that it is one in which both the fascinations and the pitfalls are densely distributed.

After the contributions on molluscs, the four on other groups (barnacles, corals, fish sagittae and polychaete jaws) seem meagre, though they are all interesting. The editors defend themselves against this criticism by stating that most work on environmental influences on skeletons is being done on molluscs. This may be true, but if so it is a shame.

A particularly good feature is the appendices. They are really instructive and will enable people, after a few mishaps no doubt, to prepare thin sections or acetate peels, or to distinguish crossed foliated structure from ordinary crossed lamellar structure. There are also useful chapters and appendices that deal with demographic and statistical matters.

One of the few unsatisfactory features of this book is that the subject of the effect of selective predation on morphology is rather neglected. It has been known for some time that the intensity of predation may affect both intraspecific and interspecific variation of shell morphology. Some of this variation is very like that produced by environmental differences; indeed it is often a real problem, even in living populations, to disentangle the two. Furthermore, though some of the differences are the result of selection taking place within a generation, some have quite high heritable components. Tevesz and Carter do mention this problem briefly in their chapter on the unionaceans, but it should perhaps have been discussed more fully somewhere.

I noticed only a few examples of laxness in editing. For instance, to give what is essentially the same equation as

$$L_t = L_\infty\ (1 - \mathrm{e}^{-kt})$$

and

$$N_t = N_{\max}(1 - c\mathrm{e}^{-r_0 t})$$

in two different places is needlessly confusing. Also, in my copy (and therefore, presumably, in several hundred others) the pages in the appendices were hopelessly jumbled. This interesting, handsomely produced, well-illustrated book deserves many buyers — but they should all carefully check pages 640 onwards for continuity. □

John Currey is a Professor in the Biology Department at the University of York.

SCIENTIFIC BOOKSHOP

H.K. LEWIS can supply works in all branches of Pure and Applied Science. Catalogues on request. Please state interests.

SCIENTIFIC LENDING LIBRARY

Annual Subscription from £7.50.
(Available in U.K. only)

Reduced rates for multiple subscriptions.

Prospectus post free on request.

Quarterly List of New Books and new editions added to the Library sent post free to subscribers regularly.

H.K. LEWIS & Co. Ltd.
136 GOWER STREET,
LONDON, WC1E 6BS

Telephone: 01-387 4282
Telegrams: "Publicavit, London, WC1E 6BS."

Circle No.16 on Reader Enquiry Card.

Historians of science confront the hydra

Robert Fox

The Ferment of Knowledge: Studies in the Historiography of Eighteenth-century Science. Edited by G.S. Rousseau and Roy Porter. Pp. 500. ISBN 0-521-22599-X. (Cambridge University Press: 1980.) £25, $39.50.

OVER the past decade, the historiography of eighteenth-century science has been totally recast. In the process, change and complexity have emerged as the leading characteristics of a period which, not so many years ago, was seen as a rather pedestrian interlude between the heady days of the scientific revolution and the expansive age of modern professional science that dawned in the first half of the nineteenth century. The reassessment owes much to the 13 contributors to this absorbing volume of essays. All of them have played a major role, in areas as diverse as industrial science and technology (treated here by Donald Cardwell), the chemical revolution (Maurice Crosland), the life sciences (Jacques Roger), "mathematical cosmography" (the title of Eric Forbes's contribution) and geology (Roy Porter). In *The Ferment of Knowledge,* however, their brief was to review the state of their particular art and to relate work in their field to larger cultural perspectives. "Strong interpretations" and "coherent analysis" were the objectives of the stock-taking, and it is by these lights that the book must be judged.

What, then, are the "strong interpretations" that emerge, and where are the new beacons? Remarkably, at least one tentative answer to these questions already seems possible. This is because the most obvious historiographical casualty of the 1970s is the very one that has traditionally dominated studies of science in the age of the Enlightenment: the notion of a unified, transcendent Newtonianism guiding confrontations with nature either as an inspiration or as a target for criticism. John Heilbron's revisionism on this point emerges from his plea for a more detailed study of the practice, as opposed to the metaphysical context, of experimental physics in the eighteenth century: as he observes, the single epithet "Newtonian" is at best unhelpful when applied to a discipline which underwent complete redefinition between the 1720s and 1800 and which flourished not only in Britain but also in regions of the Continent, notably in Germany, where the writ of Newton barely ran. Simon Schaffer makes even greater play of the conflicting traditions in eighteenth-century natural philosophy: with engaging panache, he argues for an analysis that would see natural philosophy as an autonomous form of discourse informing, and being informed by, the practice of scientific investigation, without ever being part of it. In this way, he may well make better sense not only of the many-headed hydra that scholars of the Newtonian tradition have had to wrestle with but also of anti-Newtonianism of the kind proposed by John Hutchinson in the realm of matter-theory.

Some of the nice problems of interaction raised by Schaffer also appear in Steven Shapin's sophisticated study of the "social uses" of science. In his quest for a "new contextualism", Shapin eschews straightforward notions of interaction between the discrete realms of scientific theory on the one hand and social context on the other. For Shapin, the realms are inseparable, with the result that the content of scientific ideas (and not, for example, merely their diffusion) can only be understood if we also understand the social significance they possessed in the minds of their creators. Such an approach has already yielded a fundamental reinterpretation of Boyle's Christianized corpuscularianism (in the work of J.R. Jacob), and a consideration of differing social uses may also be the key to one of the great conundrums of eighteenth-century historiography: the existence of what appear to have been competing Enlightenments, including both "radical" (the term is Margaret Jacob's) and High Church versions which repudiated Newton's carefully articulated theism.

The Ferment of Knowledge would have been welcome enough if it had merely reviewed a fruitful field of research; the editors and their well-chosen contributors have ensured that it fulfils that task admirably. But the book also says much, by implication, about the state of the history of science as a whole. Time and again, it demonstrates the unprecedented receptiveness of historians of science to work in other disciplines. The techniques of anthropology and modern social history make a welcome appearance (the latter being used to particularly good effect in W.F. Bynum's richly documented paper on medical care). Michael Foucault, too, flits across the stage, though to responses that range from the critically tolerant to the dismissive. The marriage between the history of science and the social sciences would appear to have been fully consummated, to the obvious benefit of a discipline which has striven long and hard to divest itself of an embarrassingly antiquarian past. The aptness of the term "ferment" when applied to recent developments in the history of science needs no underlining. □

Robert Fox is Reader in the History of Science at the University of Lancaster and President of the British Society for the History of Science. He is the author of a critical edition of Sadi Carnot's Réflexions sur la puissance motrice du feu *(Vrin, Paris: 1978).*

nature

28 May 1981

15\2\82

Does creation deserve equal time?

Reports from Arkansas (of which there are not many) suggest that the creationists have established a strong bridgehead there (see *Nature* 21 May, p.179). Creation and Darwin are to have equal time in the school curriculum, but how this will be done has not been decided (and may never be, if the American Council for Civil Liberties manages to overturn the bill rushed through the Arkansas legislature in March). The creationists of Arkansas would therefore be heartened by the exhibition on the "Origin of Species" mounted by the British Museum (Natural History) in London to mark its centenary this week. The exhibition, to be reviewed by Professor Barry Cox next week, is splendidly designed, persuasively instructive and, with a few exceptions, philosophically unobjectionable. But the exceptions are significant and striking. At the entrance to the exhibition gallery are two panels, one of which accurately describes the origins of Darwin's theory of the "Origin of Species". The other panel, almost but not quite as prominent, says flatly that there are other ways of accounting for the diversity of species, among which "the theory of Creation" stands out. The doctrine of equal time, hitherto used for deciding which politicians should speak on television for how long during election campaigns, is clearly making headway. Should it?

What can creationists fairly claim of time in the school curriculum or of space in public exhibitions? To what extent should teachers and exhibition designers pretend agnosticism when explaining neo-Darwinian notions of evolution by natural selection because they know that many of those they are addressing sincerely hold conflicting beliefs? Superficially, these questions resemble that of whether Darwin's theory, as now elaborated, is a scientific theory in Sir Karl Popper's sense (see *Nature* 12 March, p. 75 and later *Correspondence*). Certainly there are connections. If Darwinism is untestable by Popper's criteria, and thus by his narrow definition metaphysical, and if Creationism inspired by religious belief is also (but more frankly) metaphysical, are not the two theories equipotent and thus deserving of equal time? These are the arguments that have enabled those who call themselves scientific creationists to make such headway in the past few years, in the United States in particular. It is also likely that many of those who urge that Darwinism is "only a theory" do so because it does not conform with their religious convictions. Yet the issue of equal time is an issue in its own right, and one that deserves a simple resolution.

The issue is not a simple extension into this century of the fierce rows about God and Darwinism that occupied some of the Victorians, especially in Britain. Then, with religion firmly entrenched and sometimes backed by the force of law, it was natural that people such as Huxley should have seized on Darwinism as a way of substantiating their long-standing disbelief. Now, with the shoe on the other foot — religion less widely practised and in most places disestablished — the argument is different and more subtle. The scientific community is not, as popular legend has it, monolithically irreligious even if card-carrying believers are proportionately fewer than in the general population. And the fact that science and religion can coexist is a sign that, behaviourally, science and religion are not always antithetical. People, it seems, can spend their weekdays in laboratories observing phenomena which they suppose are causally determined and Sundays on their knees, believing that some events are differently arranged. (Sects differ in their estimates of the importance of physically non-causal events, as evidenced by their differing assessments of the function of prayer somewhere between the extremes of cathartic wish-fulfillment and putatively effective interruptions of causal processes.)

To many professional people, this awkward philosophical compromise is offensive, but mistakenly. It is not even a philosophically untenable position to suppose that the present world is well-determined, but that it was all set in train in, say, 4004 BC, the year that Bishop Ussher figured out that the world began. Continuing inspired non-causalism is harder to accommodate, but even that circle can be squared in the minds of many people, professional scientists included. But much of modern religion is drawn in such a way that inconsistency crops up less openly. In any case, in most fields of scientific enquiry, the supposed conflict between science and religion is entirely irrelevant. For the practice of science, what matters is that investigators should be unprejudiced in their expectations of the outcome of an experiment or a series of observations, and unbiased in their interpretation thereof. Even so, with the passage of time, once inexplicable phenomena have become susceptible to physically causal explanations, so that the number of unsolved problems behind which God might be lurking has shrunk. And there are some fields of enquiry in which some (not all) religious convictions are an impediment to scientific enquiry. Fundamentalists who become palaeoanthropologists have a hard life, as may devotees of astrology who work in stellar dynamics. For the rest, however, the scientific community has a duty to be as tolerant of religious conviction among its members as of, say, differing political convictions. For the most part, it is.

The trouble, in Arkansas and the British Museum (Natural History) — which should change its name — is that tolerance has gone too far. It is one thing to be tolerant of religious conviction and quite another to give religiously inspired hypotheses equal time with theories which are scientific in the sense that (correct or otherwise) they purport to relate to effects and causes which are themselves capable of being understood. Correspondence in the past few months has shown that it is possible to argue, within the convention that effects are related to identifiable causes, about the validity of neo-Darwinism. The theory of Creation is not, as the museum's placard suggests, one of the possible refinements of neo-Darwinism that may in due course emerge from more careful study of evolutionary data but a radical alternative which has the further property of not being a scientific theory: however stated, the theory of Creation relates effects to causes which are *ad hoc* and, by definition, not susceptible to understanding.

If the museum had sought to advertise to those who will be flocking to its new exhibition during the next few months that the "theory of evolution would be abandoned tomorrow if a better, theory appeared" (*Nature* 12 March, p.75), it would have done better to refer as an alternative to the theory that Hoyle and Wickramasinghe have been advocating in the past few years, which entails the repeated replenishment of the surface of the Earth with life forms from elsewhere. That, at least, will be capable of verification in the years ahead. If the objective were merely to avoid offending the small proportion of fundamentalists among its religious visitors, it could have simply said that Darwin's theory is the most persuasive of the scientific theories of evolution so far, and acknowledge that some people do not look for a scientific explanation of everything. To refer to the theory of Creation as an alternative to Darwinism is, however, to fudge the issue. Creation is not a different scientific theory but an explanation of the diversity of species of a different kind. Science museums, and science teachers, have the task of

0028-0836/81/220271-02$01.00

cultivating an understanding of what scientific explanations are. Naturally, if some among their audiences prefer supernatural explanations, toleration is in order. But not equal time.

Organizing change

Last week was formative, even traumatic, for the British government's arrangements for the sponsorship of technical innovation. It began with a report from the Advisory Council for Applied Research and Development on operations of the National Research Development Corporation, set up immediately after the Second World War as a means of exploiting inventions not taken up by private industry, and which has been for the past two decades the chief means of exploiting inventions arising within publicly supported laboratories, in universities and elsewhere. The week finished with public confirmation that the corporation is to be merged, as soon as possible, with the other and more controversial arm of the British government's patronage of industrial innovation — the National Enterprise Board. By all accounts, both organizations have been working closely together since the beginning of the year, when Sir Frederick Wood, then chairman of the corporation, was appointed chairman of the board as well. Now there is to be legislation to formalize a merger.

Politically, the rearrangement will have considerable advantages. The National Enterprise Board is, for the present British government, a curious hangover from the past. It is the survivor of the organization set up in 1965 by Sir Harold Wilson's Labour government, which is best remembered for having spent a great deal of public money persuading companies to merge with each other, sometimes successfully (as with General Electric Company), sometimes unsuccessfully (British Leyland). The present British government, ideologically opposed to public intervention in private industry, has nevertheless found it necessary to keep the rump of the old organization if only as a means of disposing of the public investments that have accumulated over the years. But the government (like its supporters) has never been easy with this inherited embarrassment. It will obviously be of some assistance if the board can appear to disappear, and the name being canvassed for the new organization — the British Technology Corporation — may help in that direction.

The proposed merger nevertheless makes sense. Each of the proposed partners needs strengthening. The National Enterprise Board, accustomed as it is to persuading bankers to act in unfamiliar ways, is short of a knowledge of what technology is about. The corporation, on the other hand, knows a lot about technology but is financially unadventurous, perhaps because of its over-zealous avoidance of borrowing from the Treasury to enlarge its operations. In principle, then, the impending merger should provide an organization more effective than the sum of its parts. The combined organization may become the elusive ideal after which British governments have striven for decades — a cost-effective way of using public money to encourage industrial innovation.

This is why all those concerned with the merger should read the report of the advisory council with close attention. It is possible that the problem of supporting innovation with public funds has been transformed since the corporation and the board were first conceived. One of the complaints that the council makes against the corporation is that it has in the past been too concerned that innovations coming its way should already be protected by patents. Somebody convinced that his new mousetrap will drive all other mousetraps from the market must first take out a patent and then, if his invention has not been snapped up by a mousetrap manufacturer, must ask the corporation to help with the cost of development. The trouble, the council properly says, is that in present circumstances many of the most interesting innovations consist not of patentable inventions but of know-how, "intellectual property". The complaint against the National Research Development Corporation is indeed softened by a recognition that the corporation has recently been changing its ways, but plainly a merger with the board will help in this direction, if only because the board is more used to backing companies than patents.

Another of the council's complaints against the present arrangements bears on the relationship between the National Research Development Corporation and the universities and government laboratories. Hitherto, the corporation has had a right to first refusal of all innovations springing from the use of public funds. It has not always been imaginative in its dealings with innovators in these places. Rightly or wrongly, many in this constituency have a sense of grievance. The advisory council now suggests that this monopoly should be dispensed with, on the grounds that there should be more than one source of support for promising ideas. This argument is not accepted by the National Research Development Corporation (which claims already to have acted on the council's other recommendations). The reluctance is understandable but unjustifiable.

Too much TV?

The geosynchronous orbit is a scarce natural resource. There is only one orbit about the Earth in which objects will appear to be stationary from any point on the surface of the Earth. Obviously such a unique facility should be used sensibly. By good fortune, national ambitions to put satellites into this orbit are moderated by international negotiation and agreement within the International Telecommunications Union, originally established to see that terrestrial broadcasting systems did not interfere with each other. Aware of the technical possibility that geosynchronous satellites might be used for broadcasting television signals directly into people's homes (with the help of a small dish antenna), the union called an international conference in 1977 to carve up the geosynchronous orbit among its member states. The report of the British Home Office (which combines responsibility for the regulation of domestic broadcasting with that for such matters as the administration of British prisons) shows that the decisions then taken should be reconsidered.

The report (see page 277) is stimulated by the British Government's wish not to waste the five channels of television broadcasting doled out to in 1977. But this is an awkward time. The Home Office has just suffered the trauma of deciding that British television-watchers should be allowed a fourth channel, to be run under the supervision of the Independent Broadcasting Authority and financed from the profits of the commercial television companies who occupy one of the three existing channels of domestic television. Much of the Home Office's strategy on broadcasting in the past few months has been designed not to harm the profits of the commercial television companies, for that would jeopardize the "success" of the fourth channel. Only this can explain the caution with which would-be operators of cable networks are being licensed. The offer of a further five television channels is not a gift but an embarrassment.

The British Government deserves some sympathy in its embarrassment. The 1977 decision that each member of the International Telecommunications Union should be entitled to five geosynchronous direct-broadcasting channels was ill considered and premature. Then, as now, it was not clear what balance that will eventually be struck between direct broadcasting to people's homes and the more efficient use of the geosynchronous orbit by means of broadcasts to community antennas. The American members of the union had the sense to opt out of the 1977 decision, and will consider how the geosynchronous orbit should be carved up among themselves only in 1983. Elsewhere, the mere availability of the new channels is certain to lead to a waste of effort in the development of broadcasting system whose economic place in the market cannot be determined while national governments (the British included) decline to let competing and complementary systems, such as cable networks, develop freely.

0028-0836/81/220272-01$01.00

Budget cuts cast shadows overseas

US agencies count costs of Reagan

Washington

The Reagan Administration, stung by criticism from Western allies of budget-based decisions to terminate or substantially withdraw from a number of international scientific projects, is taking a close look at ways in which it may be able to repair some of the damage and prevent further unnecessary friction in the future.

Of particular concern to the new Administration is that the foreign policy function of certain types of international scientific agreements — for example, a science and technology programme agreed with Spain in 1976 as part of an exchange for being allowed to place US military bases in the country — could be jeopardized if the projects are evaluated merely on the strength of their scientific merits and the vocal energy of their domestic constituency, often very small within the scientific community.

Already, Secretary of State Alexander Haig has written to Mr Reagan's budget director, David Stockman, complaining of the fact that cuts imposed either directly by the Office of Management and Budget (OMB), or indirectly by internal decision within agencies such as the Department of Energy or the National Science Foundation (NSF), have been carried out in a manner which constitutes a unilateral abrogation of international commitments.

Among the projects listed by Mr Haig which appear to have been treated in this way is the international solar polar mission, planned jointly with the European Space Agency, from which the National Aeronautics and Space Administration (NASA) has proposed withdrawing its planned spacecraft (a decision which both congressional committees and NASA officials are busy trying to reverse, perhaps through a reduced NASA commitment which would involve the European Space Agency building the two spacecraft involved). Another is the possible cancellation of national energy assessments supported by the Department of Energy and already promised by embassies in countries such as Greece, Tunisia and Venezuela.

What concerns the State Department most is that the scientific attachés of foreign embassies in Washington are telling their capitals that US promises of scientific and technical collaboration should be looked at sceptically in the future. "While budget reductions are a clear goal for this Administration, one of its principal foreign policy objectives is to render the United States a reliable international partner," wrote Mr Haig — who had previously interceded with OMB to rescue foreign aid funds and NASA's Galileo mission to Jupiter, planned with heavy West German involvement — in his letter to Mr Stockman.

International collaboration in scientific programmes is seen by Washington science policy officials as falling into three categories: that carried out by individual scientists and their institutions as part of the normal process of science; that in which both sides receive the benefit of more cost-effective technical knowledge; and that carried out with some broader foreign policy goal in mind.

It is the last of these three which is particularly threatened, particularly when an international treaty or bilateral agreement has been arranged through the State Department, and then handed to another agency for execution.

One major dilemma now facing the Reagan Administration, for example, is what to do about the International Institute for Applied Systems Analysis (IIASA) in Vienna. Faced with the need to make cuts of over $300 million — almost 25 per cent — in its proposed 1982 budget by the new Administration, NSF decided to eliminate the $3 million US contribution to IIASA for which it is at present responsible.

Unless this decision is changed by the Administration, IIASA would have to close in 1983. Scientists claim that the institute has been carrying out some useful research, even though some express doubts about its full value. However, IIASA also provides an important channel for East/West communication.

Moves are therefore under way to discover whether the closure of the institute — which would anger the Austrian government considerably, since it has put a lot of money into providing facilities — can be averted. For the 1982 contribution, already committed under IIASA's constitution, NSF is working out whether it can provide money from other sources within the foundation.

Recognizing that the problems raised by the controversy over the international solar polar mission and IIASA contributions have deep roots in beliefs about the proper political role of the federal government in support for science, the State Department has set up an inter-agency committee to discuss possible guidelines for the future. Two suggestions it is likely to discuss are that OMB should be presented, early in its budget cycle, with an overall picture of overseas implications. The other is that the State Department itself might be given funds for supporting international scientific activities which it feels have important political functions, but might not generate the required support within an individual agency.

At the same time the agencies themselves are looking closely at their own policies and procedures. Few solutions are in sight, but one consolation to US scientists is that, as a result of recent events, the whole issue has been placed high on the State Department's agenda at an early stage and is already receiving close attention from what are usually described as the "top levels" of government. **David Dickson**

Anxieties of Oslo secrets trial

Stockholm

A trial whose outcome could affect freedom of research is being held in Oslo. Two defendants are at present accused of revealing information prejudicial to Norway's security. But this is not a normal spy case: the defendants collected the information exclusively from public sources as part of a research project funded by the Norwegian Research Council for Science and the Humanities, and there is no suggestion that they intended to pass it on to any foreign power.

The two men are Nils Petter Gleditsch of the Oslo International Peace Research Institute (PRIO) and Owen Wilkes, a New Zealander formerly at PRIO and now at the Stockholm International Peace Research Institute. They maintain that freedom of research should be upheld, and that they have not in any case revealed information which could damage Norway's security. They say that what they have discovered from open sources, a foreign power (that is, the Soviet Union) could investigate in far more detail with intelligence satellites. They conclude that the Norwegian government's reticence about military affairs is keeping secrets from the Norwegians themselves, not from the Russians.

The history of the case is this: Gleditsch and Wilkes published their research report, "Intelligence installations in Norway: their number, location, function and legality" in February 1979 as part of a PRIO project (still continuing) on the location and functions of military facilities in Norway. They timed the publication to coincide with another trial then being held, in which three Norwegians — a publishing house executive and two journalists who had collected the names of Norwegian secret servicemen — were accused of collecting information which could damage the country's national security. Gleditsch and Wilkes wanted to show how easy it is to obtain from open sources information considered to be very sensitive.

The prosecutor general ordered an investigation into the Gleditsch–Wilkes report, and in March 1979, the chief of

0028-0836/81/220273-02$01.00

defence staff, General Sverre Hamre, told the Norwegian police that he considered it a threat to the activities of Norwegian defence and national security. Having questioned the authors and PRIO staff closely associated with preparing the report, the police completed their investigations in June 1979 and sent preliminary charges against Gleditsch to the court. They charged him on three counts: two under the penal code — revealing information which ought to be kept secret in the interests of national security relative to foreign powers, and acquiring this information or making it available to others — and one under the law on defence secrets — recording, copying, or publishing sketches of fortifications or associated installations. According to Norwegian practice, the police also appointed three independent experts to evaluate the report. The intelligence service proposed two of these (one was the former head of the service) and Gleditsch and Wilkes the third. Not surprisingly, the experts nominated by the intelligence service supported General Hamre's opinion that the report threatened national security. The formal indictment charges Wilkes as well as Gleditsch with all three charges, which carry a maximum penalty of 4½ years imprisonment.

The technical intelligence facilities described in the Gleditsch–Wilkes report collect and analyse data from electronic intelligence satellites, monitor and analyse military and diplomatic radio signals, and intercept and interpret electronic and telemetric communications. Although the authorities are traditionally tight-lipped about such installations, the authors point out that it was very easy to find out about them by using, for example, the ordinary telephone directory and public lists of government-owned property. They also learned a lot by observing the installation's antennas: their size, shape and layout. The authors maintain that if the Soviet Union, with its sophisticated spy satellites and, presumably, other more traditional techniques of spying, is interested in the installations, it would already know a great deal about Norwegian security. They claim that the authorities' secrecy is mainly directed towards the Norwegian population itself, pointing to police reaction to the publication of their report in Norwegian, shortly before the trial began. For two years the report has been available in English and the police have made no attempt to confiscate it; however, they have now declared the publication of the Norwegian translation illegal.

Support for Gleditsch and Wilkes has come from the International Peace Research Association, the Norwegian Writers' Union and political youth organizations in Norway. The trial is expected to last for another week or so, and the verdict will follow a couple of weeks after that. One or both sides will almost certainly appeal. **Wendy Barnaby**

UK nuclear power

Plans panned

For the second time this year, Britain's Central Electricity Generating Board (CEGB) has received a rap over the knuckles for not providing enough information about the economic case for its planned nuclear power programme. Hard on the heels of sharp criticism from the House of Commons Select Committee on Energy last March, the Monopolies and Mergers Commission last week published a report on the board's finances (HC 315, HMSO, £9.30). The Monopolies Commission was asked to find out whether CEGB is charging a fair price for the electricity which it supplies to the consumer through regional electricity boards.

The commission says that one of the chief reasons for the high cost of electricity in Britain is CEGB's previous tendency to invest in new plant before strictly necessary. Decisions to order new plant, it concluded, have often been based on over-estimates of future demand, and economic cases have been made on over-optimistic estimates of construction times, capital costs and the ultimate performance of power stations.

The commission's report is particularly critical of the decision, to build the second advanced gas-cooled reactor at Heysham, taken in 1979 as part of the government's decision to embark on a further nuclear programme, probably based on the Westinghouse design of the pressurized water reactor. Failing approval at next year's public inquiry, CEGB hopes to keep open the option of basing the programme on the notoriously costly British-designed AGR by building Heysham II.

Although the commission acknowledges that CEGB has learnt from its previous mistakes with the AGR, it still criticizes the board for not providing sufficiently detailed cost estimates for Heysham II, even though the decision to build it was taken largely on strategic grounds.

The commission also faults CEGB's case for future nuclear power stations on economic grounds. It says that the board's estimate of net effective cost assumes unrealistic improvements in future construction time and operational performance. Margins of error in cost estimates, together with a better indication of the board's assumptions, should have been presented to the government when making an economic case for the nuclear programme. The board should also have compared the cost of building new nuclear stations with the cost of refurbishing old coal-fired stations.

The commission, however, has praised the way in which CEGB keeps a check on its current expenditure and seems to be impressed with the performance of its Barnwood division, unpopular with many of its contractors, in keeping a detailed check on cost increases during construction. The commission's report, however, ends on a rather gloomy note. Even if CEGB implements all its recommendations, the commission cannot foresee any substantial cut in electricity prices to the consumer. **Judy Redfearn**

Gloom on research

A high-level meeting of research councils from Germany, France, the United Kingdom and the United States was held in the English country town of Abingdon on the weekend of 16–17 May. The occasion, arranged by Sir Geoffrey Allen, chairman of the British Science and Engineering Research Council, is the second in a series of consultations begun with a meeting called by Professor H. R. Leibnitz, then the president of the Deutsche Forschung Gemeinschaft in June last year.

The objective appears to have been an exchange of views, most of them gloomy. The United States participants (Dr John Slaughter and Dr H. R. Langenberg, respectively director and deputy director of the National Science Foundation) left their fellow-administrators with the impression that the Reagan budget had been a body blow whose effects will not be confined to the United States, but that the United States is also worried by more long-standing problems — the difficulty of recruiting teachers of engineering for United States universities, for example.

European participants in the meeting are said to have welcomed their transatlantic colleagues into the company of the impoverished, and to have urged that if the United States now faces a period of entrenchment it would be worthwhile thinking of collaboration with European collaborative ventures such as those at CERN. The occasion seems also to have been one for concerted European complaint about the budgetary threat to the solar polar mission, the plan to send two spacecraft (one European, one American) into polar orbits about the Sun.

US science funding

Rewards of genius

Washington

Acting on the theory that intellectual and cultural breakthroughs can be accelerated if latent genius is permitted to flourish free of the more material considerations of making a daily living, one of the newest — and wealthiest — foundations in America has announced the first of a series of five-year, no-strings-attached awards to "exceptionally talented" individuals.

The awards are being made by the trustees of the John D. and Catherine MacArthur Foundation set up two years ago from the estate of the late Chicago insurance and real estate millionaire. The

0028-0836/81/220274-02$01.00

foundation, which has assets of $841 million, last week announced 21 individual awards worth a total of over $4 million, with another 29 expected later in the year.

Eight of the initial recipients are scientists, and include Stephen Wolfram, a 21-year-old physicist at the California Institute of Technology, the Harvard geologist and paleaontologist Stephen Jay Gould, and oceanographer and climatologist John Imbrie of Brown University and the University of Rhode Island.

The names were chosen by the trustees of the foundation from a list of nominations proposed by 100 educators, scientists and artists who acted as scouts in what other foundations, more conventional in their granting of awards, have dubbed the "search for genius". There are no conditions attached to the way that the money — between $24,000 and $60,000 a year for each individual, depending solely on age — can be used, nor can the award be withdrawn within the five-year period.

The MacArthur Foundation's novel approach to the support of intellectual activities is perhaps the most ambitious of a number of attempts to meet the charge that more conventional forms of funding discourage innovative or risk-taking work.

Last year, for example, Berkeley physicist Richard A. Muller, winner of the National Science Foundation's (NSF) Alan T. Waterman award based on research for which he had initially encountered difficulty in obtaining support, told a congressional committee that individuals engaged in innovative research often had similar experiences, for example when their proposed project did not fit neatly into one or another disciplinary compartment.

Dr Muller's testimony and other similar complaints led Congress to ask NSF to assess current funding mechanisms to find out how well they are working. So far these studies have not uncovered any substantial problems, nor pointed to any particularly radical solutions.

Furthermore, a task group set up last year by NSF's advisory council on the funding of innovative high risk proposals has reported that "on the whole, the foundation's procedures seem to be effective".

In the light of the comments received and of its own investigations, the task force, headed by Halsey Royden, dean of the school of humanities and sciences at Stanford University, suggested that NSF programme officers be given greater encouragement to support innovative risk-taking proposals and that a small Group on Innovative Research Topics be set up under the deputy director, to "promote promising research that does not fit naturally into the framework of existing programs and divisions of the NSF."

Both proposals are now being considered by NSF. However, Dr Langenberg points out that the task force specific recommendations are likely to be absorbed into the bigger organizational changes now under way. **David Dickson**

High-energy physics

On the rocks

The Swiss tunnel expert Giovanni Lombardi, who has honeycombed the Alps with road and rail tunnels, denied last week that it might be impossible to build the tunnel for LEP — the next big project of the European nuclear physics laboratory CERN.

The assertion, in the British magazine *Consulting Engineer*, would have prevented governments from approving LEP construction at next month's crucial CERN Council meeting. Lombardi is not only a world-recognized expert on Alpine tunnelling but also CERN's principal geological adviser.

Lombardi and CERN do however admit to geological difficulties in that part of the LEP tunnel which will go under the limestone of the Jura, to the north-east of the CERN site. The worst of the troubles have however been avoided by shrinking LEP from 30 km to 27 km circumference, it is said. In this way, the tunnel should avoid the folded core of the Jura, a region of unstable limestone which Lombardi knows to contain water and mud-filled caverns.

Even so, progress through the Jura will be uncertain and based on "forage à l'avancement", where a small (2-inch) hole is drilled 20–30 m ahead of the main borer to probe for boundaries between limestone layers. At hundreds of metres below the water table, there may be mud and pebble-filled "karsts" at these rock divisions which will have to be emptied and filled with concrete before drilling through. If water-flow through the karst is too great, it can be impossible to place the concrete.

Uncertainties of this kind have persuaded smaller member states of CERN, led by Sweden, to demand guarantees that the CERN budget will not be raised to meet any extra costs. It has thus been agreed that future CERN budget increases can be vetoed by any state, while for decreases a two-thirds majority is sufficient. And CERN's director-general, Herwig Schopper, has agreed in principle that LEP cost escalations would be met by lengthening the time over which LEP is built.

Strenuous efforts to delineate the geology of the Jura are under way at CERN, but the principal reconnaissance gallery will not reach the tricky region until April next year. So CERN is also drilling a hole vertically above the deepest part of the proposed tunnel under the Jura and making geophysical observations from within it to find the water table and the trend lines of the various limestone boundaries. The results of this investigation will not, however, be known for another three months.

Meanwhile CERN is preparing for its mid-year council meeting on 25 June at which delegations from the 12 member states would normally approve the 1982 budget. This year the budget contains an appropriation for LEP, which is not being costed separately. The debate will centre on precisely what level the budget should take, and what guarantees can be given on LEP cost overruns.

Sweden, apart from its doubts on the latter score, is in political crisis, and will almost certainly abstain; Norway may do the same; and the Netherlands are in the midst of elections and cannot predict their position. Moreover, the CERN Council delegations of many of the member states have not yet been officially briefed (this includes Britain) and so the outcome is far from certain. Procedurally, if eight states vote for the budget including LEP, and none votes against, LEP can go ahead, and this seems likely, although Schopper would like to leave the door open for a few months after June to achieve a unanimous decision. **Robert Walgate**

Hormone legislation

Consumer protest

Brussels

The failure of the European Community's council of agriculture ministers to make significant progress on banning the use of natural and artificial hormones in livestock production is forcing European consumers to take retaliatory measures. The Bureau of European Consumers' Association is now trying to persuade the sympathetic member states to block meat imports from the United States, New Zealand, Australia and elsewhere.

Last September, the Community agreed in principle to ban the use of all hormones in livestock breeding. The decision was hailed as a victory for the consumers but has since proved to be a hollow one. It has been suggested that the September council failed to understand the difference between natural and artificial hormones and hence the problems of forbidding the use of the former. On 12 May, the agriculture ministers met to consider the European Commission's two proposals for directives to implement the ban — the outcome was disappointing.

A German proposal was adopted banning some artificial hormones already forbidden under existing laws operating in all member states except the United Kingdom. Diethystilboestrol and other stilbenes are now to be banned, although whether this entails a separate directive or merely the partial implementation of the Commission's all-embracing directive is unclear. The other growth hormones, and the problem of enforcing any bans, will again be considered by the next agriculture council on 15 June.

The United Kingdom is becoming increasingly isolated in the discussions. The philosophy of not to ban a hormone until it has proved to be dangerous resembles that of the United States, but the legislation of other Community countries reveals a much

0028-0836/81/220275-02$01.00

more cautious attitude. The United Kingdom has, however, been able to persuade its partners not to adopt the wide-sweeping legislation the Commission proposes until further studies have been carried out.

Various scientific committees will now be given the job of examining the health hazards of natural oestradiol, testosterone, progesterone, trenbolone (a processed natural hormone), zeranol and perhaps others. The United Kingdom rules permit the use of these hormones, which many other member states have forbidden. Similarly, until there is "conclusive" scientific evidence to the contrary, oestrogens, androgens and gestagens (except stilbenes) are considered safe, as far as the Community is concerned, for therapeutic use and to regulate the menstrual cycle.

The consumers feel that this scientific research could drag on for a long time, and matters will not be helped by the United Kingdom taking over the council presidency after the summer from the Dutch, who support tough Community legislation. **Jasper Becker**

Canadian radioastronomy

Longer bases

Washington

Canadian astronomers have started design work on what it is hoped will eventually to be the largest radioastronomy facility in the world. An eight-dish array of 32-metre antennas is to be distributed along a 5,000-kilometre line stretching from one end of the country to the other.

Canada's National Research Council (NRC) has already given its formal approval to the project, provisionally named CASCADE (after the Canadian Astronomy Society). Meeting in Ottawa in February, the council members gave the array top priority over four other possible national facilities that had been suggested for funding: a high-energy electron ring, an orbiting observatory, a kaon antiproton physics facility, and the creation of several centres of acoustics.

NRC and the Natural Sciences and Engineering Research Council have now agreed to share the costs of a $300,000 design study for the array being carried out by a technical committee chaired by Dr Ernest Seaquist of the University of Toronto's department of astronomy. It is hoped that detailed plans for the array, estimated to cost about $30 million at 1981 prices, can be completed by the beginning of next year; and that if the Canadian government can then be persuaded to support the project, funding for construction be provided in 1983.

Plans for a very long base array — which will be able to provide an angular resolution of 5×10^{-4} arc seconds at a wavelength of 1.5 centimetres, a hundred times better than can at present be provided by the largest Earth-based radiotelescope — have been discussed for the past three years within the Canadian astronomy community, one of the first to develop and work with very long baseline interferometry. Two years ago, the very long base array came out on top of two other proposals, a 100-metre centimetre-wavelength dish and a large (25–30 metre) millimetre-wavelength antenna, in a study conducted by the Canadian Astronomy Society.

Despite enthusiastic support from astronomers, there was initially some scepticism from industrialists in NRC, who doubted that the project would have sufficient economic pay-off to justify the initial capital investment. However, after considerable lobbying several companies, particularly in the electronics, computing and telecommunications industries, were persuaded to back the proposal, particularly after it was pointed out that most of the components would be built in Canada. Astronomers argued, for example, that the increased production of antennas and equipment, together with some improvement in high frequency performance, would help Canadian aerospace and communications companies to compete for foreign markets.

As planned, the array would stretch along an east–west line from Newfoundland to Vancouver Island. Signals would be recorded independently at each station and subsequently correlated at a central processing facility, giving 28 components of the Fourier transform of the image of a cosmic radio source.

Efforts are being made to generate public support for the array by appealing to nationalistic instincts. A circular distributed by the chairman of NRC, Dr Larkin Kirwan, says the array would "assure Canadian leadership in radioastronomy for at least thirty years", and the Canadian Astronomy Society says it would "remain a permanent advertisement for Canadian science and technology".

In a similar vein, the secretary of the design committee, Dr Brian Andrew of NRC's Herzberg Institute of Astrophysics, heads a public relations committee whose task he describes as being "to elevate the array to the status of a national shrine." He also points out that the planned array would be able to take advantage of Canada's unique geography, and that "linking the country from end to end seems to have intangible overtones that would make it politically attractive".

Meanwhile budget restrictions have forced US scientists to put back their own plans, developed last year by a group at the California Institute for Technology and its Jet Propulsion Laboratory, for a two-dimensional array stretching across the continent of the United States, and including antennas on both Hawaii and Alaska (*Nature* **288**, 4; 1980).

Similar in size, conception and cost to the Canadian plans, the US proposal is said to have been given top priority for the next decade by the Field committee now preparing a report on the future of US astronomy for the National Academy of Sciences. A north–south array would fit well with Canada's plan.

However, one of the victims of the budget cuts announced by President Ronald Reagan in March was a 25-metre millimetre-wavelength radiotelescope planned for Mauna Kea in Hawaii, and previously approved for funding through the National Science Foundation by the Carter Administration. With this project placed back in the melting pot it seems unlikely that other new capital construction will get much consideration in the next few years. **David Dickson**

Hungarian agriculture

Economic growth

Hungarian agriculture is to be remodelled in the next twenty years on ecological principles. This is the burden of a report last month to the Academy of Sciences of an interdisciplinary survey of the country's "agroecological potential". The survey was first proposed at the 1978 annual general meeting of the academy by Dr Istvan Lang, at that time deputy general secretary. Thirty research institutes, universities and computer centres took part, and more than 400 scientists were involved.

The survey asked three main questions: what quantity of agricultural plant production can realistically be attained by the end of the century? What conclusions can be made about the long-term targets of economic policy? How can production be increased and costs reduced in the medium term? The preliminary conclusion is that in the most favourable conditions, the annual grain yield could reach 22 million tonnes (present level 12–13 million tonnes) and that the productivity of grasslands could be doubled.

Nevertheless, the survey notes, there are constraints on the development of Hungarian agriculture. The country is poor in fossil fuels, so that with rising oil prices a point may be reached where it is economic to settle for lower than maximum yields and less expenditure on fertilizers. The country's geology is also a constraint. Hungary lies in the lowest part of the Carpathian basin, and the run-off from the mountains has produced tracts of saline and/or alkaline soils, principally in the western half of the country. Thus the Puszta remains a virtual desert in spite of an annual precipitation of 550 mm. Yet normal leaching methods of reclamation are ineffectual, since the soil is so impermeable that the water would simply pond on the surface. Moreover, Hungary has no means of disposing of the drain water, since the quality of the rivers Tisza and Danube must be ensured at the southern frontier.

Nevertheless, on the Comecon scale of

"relative climatic productivity", which assigns an index of 100 to the Soviet Union south of latitude 60°, Hungary scores 139, second only to Bulgaria with 145, and agricultural products, especially fruit and vegetables, account for some 30 per cent of Hungary's export earnings.

The goal of relating agriculture more closely to ecology was accepted in the five-year plan approved last year. The Hungarian economy has, however, come a long way from the central planning of the post-war Rakosci regime which, in an attempt to make the country self-sufficient, ordered certain collective farms to grow cotton and citrus fruits — with predictably poor results. Since the 1968 economic reform, the managements of the various collective farms have complete autonomy in deciding what they should grow. The only constraints are now those imposed by market prices and government incentives.

Often the consequences have been nonsensical. In the great sugar-beet affair of the mid-1970s, farmers applied more fertilizer than recommended and produced gigantic beet with no increase in sugar content. (Bonuses are now related to sugar yield, not gross weight.) In a more recent scandal, farms have obtained land improvement grants for top-soil dressing and then applied the soil to areas which were already of fairly high quality. This produced money for the farms, and excellent statistics for the soil improvement service — but a net loss to the economy.

The next step towards ecologically orientated agriculture will depend on the working out of suitable incentives for collective farms — free, it is hoped, from the loopholes and misunderstandings of the past. **Vera Rich**

Direct satellite broadcasting

More TV ahead

Britain could start direct broadcasting by satellite in the mid-1980s. That is the government's response to the report of a Home Office study commissioned a year ago and published last week. The government favours a modest start in about 1986 beginning with two television channels and possibly some information services.

No change in the way broadcasting is controlled is, however, proposed, nor will the British government, unlike its French and German counterparts, help finance the venture. It hopes that British industry and the broadcasting authorities will be able to raise the money, estimated by the report to be £75–£160 million of capital to establish a system, depending on the type of satellite and the number of channels chosen. Throughout a 10-year period, the annual cost of operating a channel, including original capital costs, is estimated at £10–16 million. In addition, the broadcasting authorities would need to spend between £10 and £100 million a year to run the service for 50 hours a week.

The Home Office study was stimulated by the fear that delay could put British industry at a disadvantage in the supply of equipment to what is expected to be a rapidly growing world market. The Home Office also had to work out a way of using the allocation in 1977 of five television channels from geostationary satellites.

There are two possible systems to choose from — a modified version of the European Communications Satellite and a satellite based on L-sat, the prototype of the second generation of telecommunications satellites due for approval by the European Space Agency next month. Signals could be beamed direct to individual 2–3-m dishes or to larger community antennas. Control of programmes would remain with the British Broadcasting Corporation (BBC) and Independent Broadcasting Authority (IBA), which control all British broadcasting through terrestrial systems, although coverage would extend to neighbouring countries.

The investigation uncovered a mixed response to direct broadcasting. The aerospace and electronics manufacturers are the most enthusiastic. The BBC is also keen, believing that it could use two channels — one to transmit repeats of programmes already shown on terrestrial channels and paid for from the standard television licence fee, and the other to transmit new special programmes, paid for by subscription. But the IBA, grappling with the introduction of breakfast television and a fourth television channel, does not think there will be a market before 1990.

The technical problems are less forbidding than the financial. The broadcasting authorities, which do not have £75–£160 million to invest, would prefer to lease channels from an operator. Private industry, especially the newly-formed Satellite Broadcasting Company Limited, has expressed some interest in launching a satellite and operating it as a common carrier, but industry's willingness to invest will depend on its confidence that a market can be found.

The option of using L-sat to test the British market has been preempted by the Italian broadcasting authorities, which are planning to run an experimental project.

Elsewhere in Europe, France and Germany are planning to take up their allocation of five channels by means of joint ventures. Two of the French channels will be used for broadcasting signals now transmitted terrestrially. The prospect that a third channel might be leased to Luxembourg has faded, while plans for an internationally financed Swiss satellite, aimed at providing commercial television programmes for the whole of Europe, have been set back by the rules on foreign investment in Switzerland. Similarly, the Scandinavian plan for a regional broadcasting service by satellite is in limbo.

Judy Redfearn

Cooling sackcloth

New Delhi

This year Indians are beating the summer heat by installing a novel airconditioner on the terraces of homes and offices. They spread a number of empty, used gunny bags (sack made of jute) over the terrace and keep them soaked with water round the clock with the help of sprayers like those used on lawns. This simple system works like an airconditioner. The Sun's heat evaporates the water instead of warming up the roof, and the process of evaporation cools the roof. In other words, the water-soaked gunny bags reverse the heat flow — and the heat from the house is thrown outside through the roof instead of the other way.

A fall of as much as 10°C in indoor temperatures has been observed during day times with this technique. Sleeping conditions at night are much better as the ceiling fans, a common fixture in most homes, throw out cold air because the roof is even cooler than the floor.

The "gunny bag airconditioner" has been developed by the Central Building Research Institute (CBRI), who say it is a cheap alternative to the desert coolers and airconditioners that only a few Indians can afford, and which consume electricity that is in short supply. The cost of installing the system per square metre of roof surface is 5.50 ruppees (£0.30) and the water requirement is 9 litres per day.

CBRI started experimenting with its technique on its own buildings at the Roorkee campus. Results of these experiments, carried out during three successive summers, were so good that several factories and offices became interested. The system was installed on a 1,650 square metre roof of the four-storey building of Bharat Heavy Electricals Ltd in Hardwar. After the gunny bag treatment, the roof temperature of the building fell from 45°C to 28°C and indoor temperature from 39°C to 30°C. According to CBRI scientists S. P. Jain and Vinod Kumar, "not only the top floor but also rooms in the down floor were cooled as a result of constant working of the system".

CBRI engineers say that passive cooling by roof surface evaporation is suitable not only for India but for all developing countries in the tropics where artificial airconditioning is not within the reach of common man.

So far the roof surface evaporative cooling system has proved successful only for buildings with flat roofs. Attempts to use the system in roofs that are not flat are under way, and there are plans to develop inorganic water retentive materials as a substitute for gunny bags, which need replacement every three years or so.

K.S. Jayaraman

0028-0836/81/220277-02$01.00

CORRESPONDENCE

Future reference

Sir — The recent letter from Andrew Brooks (*Nature* 7 May, p.7) suggesting that research students should take a course in information sciences recalls Dr Johnson's comment "Knowledge is of two kinds. We know a subject ourselves, or we know where we can find information about it". The combination of a continued growth in scientific literature and the worldwide cutback in library budgets is making it more difficult for us to "know where we can find information about it".

The long-term solution appears to be electronic publishing; a range of databases giving bibliographic and abstract information backed up by huge databanks where the original text is stored in machine-readable form for full retrieval. It has been estimated that in 1979 there were in the United States 93 databases in agriculture, life sciences, pure and applied science, producing 26 million records, while in Europe there are 92 databases in the same subject area producing 19 million records.

Commentators predict that these databases will grow rapidly over the next five years. Users will no longer expect to find the article they require in their library but with the aid of on-line access to the databases will be able rapidly to identify the literature relevant to their needs and to call it up from the databanks. Time spent on training should be more than compensated for by savings in time spent on the customary literature search.

Robert Campbell

Blackwell Scientific Publications,
Oxford, UK

Badger strain

Sir — In your correspondence about the Zuckerman Report, little has been said about the more medical aspects of the problem.

The report opens by confusing the issue. In the first paragraph the disease "tuberculosis" is referred to and figures given for its incidence in the United Kingdom. However, tuberculosis is a term covering the disease state produced by infections with more than one strain (some would argue more than one species) of bacterium. Man can be infected with *humanis, bovis* and *avium* strains. It is only strain *bovis* that concerns us in badgers. Paragraph 82 does rather belatedly clarify this, and implies that the figures on the first page overestimate the human risk by *sixty-fold.*

There is an even more dubious presentation of data in paragraph 58, where it is stated that the "bovine strain was four times more responsible for *abdominal* tuberculosis" in children (my emphasis). These figures are from a survey between the wars, when abdominal tuberculosis was commonly contracted by ingesting infected milk, while other forms (especially pulmonary) were transmitted more by human contact. Hence, *bovis* predominated in abdominal cases (before most milk was pasteurized), and *Humanis* in many others. (Abdominal tuberculosis used also to be common in cats, another group, besides children, that ingested large quantities of raw milk!) I fail to see why data were selected for cases in children, and for this particular anatomical site, without some explanation.

The report goes to great lengths to destroy various falsehoods that have been in circulation. The above two examples refer to the cases respectively that there is a real threat to man, and that *bovis* does infect man. The fact that they in my opinion overstate the case in the support of what is in basis a sound argument does much to undermine the authority and objective position of the report.

It is also disappointing that the chapter dealing with the measures that have been taken is so small a part of the report. Here again I think a full picture is not presented.

The report also states (paragraph 107) that vaccination is not allowed for livestock. What it does not state is that in no way can this affect the question of badger vaccination — vaccination of cattle would interfere with the tuberculin test, but the test is of no use in badgers. (Curiously, the fact that the test does not work in badgers is then used as a reason against vaccination.) No explanation is given as to why infected badgers should not be vaccinated (as would happen unwittingly if blanket vaccination were practised). There are possible objections to vaccinating infected animals, but whether these would be outweighed by the advantages of blanket vaccination is an open question.

The report does not raise the question of the effectiveness of BCG in badgers. If it is not known, I would have thought it a prime research aim. If some of the badgers could be protected for some time, this would reduce the susceptible population.

Further, long-term policy is not well explored. The report supports the ministry's policy, but does not state what effect is expected. Is the policy to eradicate the disease or just to control it? Present measures seem to be only controlling it — is not that in itself a justification for a rethink on policy?

I am convinced of the sincerity of the ministry's actions, but I am also convinced of the sincerity of many zoologists who doubt the basis of those policies. Unless those doubts are considered I do not see how there can be public confidence in the ministry's policies. The association between state veterinary medicine and wildlife epidemiology is young, but of utmost importance as long as rabies lingers on the other side of the English Channel. I for one would welcome a widely based inquiry into ministry policy on both the current problem and those plans we hope we will never need.

M.J. Chapman

Easingwold, York, UK

Paradigm change

Sir — It is now well understood that the abandonment of one time-honoured theory or paradigm in favour of another by scientists does not solely depend on proof[1]. It also takes a substantial time for the whole scientific community to accomplish such a paradigm change[2]. But for individual scientists the time taken can be much shorter. I have personally experienced several changes in my commitment to theories and paradigms, which were prompted either by external motives or self-induced discovery independent of the extant information. These changes occurred either over a period of a year or so, or, more than once, within a matter of an hour — almost like lightning.

So *one day* may be long enough for a scientist to abandon one particular theory in favour of another. In this context and in opposition to Hesketh[3], I support the scientists in the British Museum (Natural History) when they write[4]: "the theory of evolution would be abandoned tomorrow if a better theory appeared", although I admit that people would vary in the speed at which they might come to find a novel theory "better" than the present one.

In this connection, I wish to point out that the current trouble with the evolutionary theory does not reside in science itself but in the attitude of some religiously committed people. If so, science will be able to remedy the situation only when it appears trustworthy in the eye of the public. Science is *not* a religion. In contrast to followers of a religion, scientists should admit that the currently held (scientific) theory may very well be wrong[5], whatever consequences that may have for a civilization based on science and technology.

A. Sibatani

CSIRO Molecular and Cellular Biology Unit,
Sydney, New South Wales, Australia

1. Kuhn, T.S. *The Structure of Scientific Revolution* (University of Chicago Press, 1962).
2. Kuhn, T.S. *The Essential Tension*, 320–339 (University of Chicago Press, 1977).
3. Hesketh, R.V. *Nature* **290**, 286 (1981).
4. Ball, H.W. *et al. Nature* **290**, 82 (1981).
5. Sibatani, A. *Trends biochem. Sci.* **6** (in the press).

Protecting the work

Sir — Many scientists need to handle biological material in a containment facility that provides sterile conditions for the work as well as protection to the operator. This requirement highlights one aspect of the current debate concerning performance and use of microbiological safety cabinets where an increasingly held belief (based mainly on misconceptions as to the properties of particles in moving airstreams) suggests that Class I cabinets can provide considerable product protection in addition to the operator protection for which they were designed, and that consequently there is no need for the sophistication of Class II types.

We have recently carried out a series of experiments, in ordinary laboratory conditions, which clearly demonstrate that protection is not obtained in Class I cabinets whereas in Class II types, built to the requirements of BS 5726 and working correctly, there is a consistently high degree of product protection that is independent of the particulate contamination in the laboratory air.

Although the idea of using a Class I cabinet for some degree of product protection is attractive, it is evident that this is intrinsically unsound and should be abandoned.

R.P. Clark
J.P. Hill

Clinical Research Centre,
Harrow, Mddx, UK

0028-0836/81/220278-01$01.00

NEWS AND VIEWS

Cerebellar modelling

from Rodolfo R. Llinás

DAVID MARR's untimely death last year is a great loss to neuroscience; he introduced changes in this field that will remain with us for many years to come. Of the areas of work at which he excelled, two of the most prominent are his model of cerebellar function and that of the visual system. His approach tackled some of the truly central issues in neurobiology today. In particular, I refer to the problem of correlating the properties of single neurones and of neuronal nets to the global functions of the brain.

For people interested in establishing such a bridge, the cerebellum continues to be both a thorn in the side and a reason for hope. The basic dogma of the past, that from the diligent study of the electrical properties of single neurones and their resulting nerve nets would arise *ipso facto* a general understanding of brain function, without such unsavoury steps as general hypotheses, is now relegated to oblivion. This turnabout has resulted to some extent from research on the cerebellum where knowledge of its phylogeny, ontogeny, morphology, physiology and neuropharmacology, although most complete, has not *per se* provided new understanding regarding its overall function — this is the thorn in the side. Indeed, although it is fundamentally true that an understanding of brain function must be based on knowledge of the neuronal elements which constitute the brain, as the properties of the system ultimately reside in the properties of the neurones and the nets which they weave, it is also true that such knowledge must be embedded in concepts capable of dealing with the properties of the system as a whole.

One of Marr's basic contributions to neurobiology was the formulation of a global hypothesis regarding the workings of the cerebellum[1]. This formulation, which was published in 1969, may be considered one of the first general hypotheses concerning central nervous system (CNS) function to be based on the detailed electrophysiological properties of single neurones. Although other theories regarding the function of this system could be considered early models, such as that of Braitenberg and Atwood[2], they were basically derived from purely morphological parameters and, while ingenious and beautiful, were later refuted by more detailed knowledge of electrophysiology brought to light in the mid 1960s. Marr's model, in addition to proposing a broad solution, presented an interesting blend of important ideas.

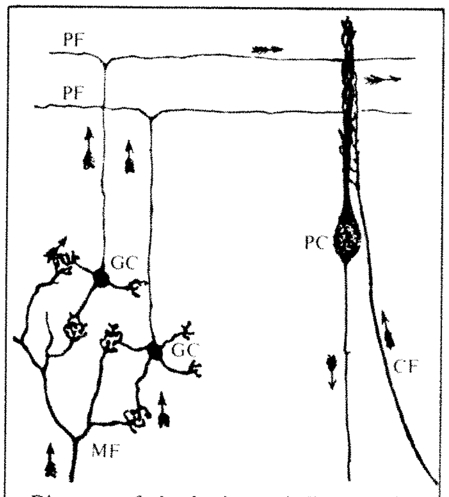


Diagram of the basic cerebellar circuit, modified from one of Ramón y Cajal's original drawings. MF, mossy fibre; GC, granule cell; PF, parallel fibre; PC, Purkinje cell; CF, climbing fibre. The arrows indicate direction of nervous conduction in the different pathways. David Marr's hypothesis assumed that activity in the climbing fibre modified the efficiency of the junctions produced between parallel fibres and Purkinje cell.

Fundamentally, the model considered the cerebellum as representing, by the activity of one of its inputs, the mossy fibres, a continuous description of the functional state of the musculoskeletal system and of other brain regions directly involved with the central programming of movement (such as the vestibular system and motor cortex). Marr's view regarding this input was quite close to the classical view in which mossy fibres from many regions of the CNS are considered the main information-gathering system which allows the cerebellar cortex to correct for possible mistakes in motor execution. The pivotal point here is that motor coordination arises at cerebellar nuclear levels by the combination of two types of nerve terminal. The first type is the collaterals of mossy fibres which wind their way to the cerebellar cortex and end directly on the neurones of the cerebellar nuclei as excitatory terminals. The second consists of the inhibitory synapses provided by the Purkinje cell axons, the only output of the cerebellar cortex. This latter input brings to the cerebellar nuclei an elaboration by the cerebellar cortex of the mossy fibre afferent information. Cerebellar coordination is then considered to arise from the 'comparison' of these two inputs.

With regard to the second cerebellar afferent system, Marr's model formally introduced the attractive idea, originally suggested by Brindley[3], that the function of the climbing fibres (which are the inferior olive's input to the Purkinje cells) is to modify the synaptic properties of the parallel fibre–Purkinje cell junctions. The postulate (that of heterosynaptic modification via a postsynaptic element) added to the concept of cerebellar control the very interesting possibility of plasticity on the basis of totally unambiguous morphology. This hypothesis is appealing because it provides a possible clue for the rather monstrous morphological differences between the climbing fibre and the mossy fibre/parallel fibre inputs to Purkinje cells. Indeed, the one-to-one relationship between a climbing fibre and its Purkinje cell (see the figure) is such that several hundred synaptic junctions are formed between the two elements establishing the most powerful chemical synapse in the CNS. This morphological arrangement contrasts vigorously with the mossy fibre/granule cell/parallel fibre system which is the richest input to any cell in the CNS with as many as 250,000 synaptic contacts, each from a different granule cell. These two inputs to Purkinje cells presented a possible example in which plastic synaptic interactions could in fact be expected. Thus, if the activation of climbing fibres could somehow modify the properties of the parallel fibre–Purkinje cell synapse in the sense of increasing the efficacy of the parallel fibre–Purkinje cell

---

Rodolfo R. Llinás is in the Department of Physiology and Biophysics, New York University Medical Center.

0028-0836/81/220279-02$01.00

junction (if the activity of the two inputs were to occur at short time intervals), this model would have gone far in solving one of the central problems in CNS function: that of the acquisition and retention of 'memory engrams'.

Unfortunately, more than ten years after its formal introduction and after much research in this area, Marr's model has fallen from favour on at least two grounds. First and most important, synaptic modification as postulated by Marr — an increased efficacy of the parallel fibre–Purkinje cell synapse by the co-activation with the climbing fibre input — has not been demonstrated. The fervour and tenacity with which this idea has continued to be tested, even without the accumulation of direct experimental evidence in its favour, speak of its seductiveness and beauty. On the other hand, the plethora of claims on indirect grounds for and against Marr's hypothesis (despite the clarity of the postulate and the unambiguous requirement for its experimental validation) is in itself quite telling, and the present decreasing number of positive claims suggests that if modifiability indeed exists, it will relate more to a decrease of the efficacy between parallel fibre and Purkinje cell after climbing fibre activation[4,5], rather than the increase in efficacy postulated originally by Marr. Although this is a major departure from Marr's hypothesis, the above claims, if correct, may be considered in accordance with the original proposal by Brindley[3]. However, direct demonstration of synaptic modifiability has yet to be shown.

Second, at another level, experimental results continue to plague the general hypothesis of cerebellar plasticity. Several lines of investigation have suggested that the cerebellum may be necessary for the acquisition and retention of organized motor responses. One such motor response is the posture and movement compensation which follows unilateral vestibular nerve lesion. This is an ideal paradigm as it involves acquisition of quite complex motor skills and vestibular nerve lesion is accompanied by a total loss of equilibrium. Moreover, because this compensation is required in real life after vestibular damage produced by infectious lesion of the middle and inner ear, it has the advantage of not being experimentally contrived. It was deemed a potential test of cerebellar plasticity on account of its being deeply impaired by cerebellar ablation in the rat. After unilateral vestibular lesion, all vertebrates demonstrate continuous rotation and rolling which would be incompatible with survival unless quickly corrected. The animals must then learn to stop unilateral rotation. Compensation is impossible if the climbing fibre afferents are removed (the climbing fibre input appeared necessary for the acquisition of vestibular compensation); so the paradigm seemed most appropriate to test Marr's hypothesis[6]. A serious problem arose, however, in testing for the retention of this compensation several weeks after the surgical section of one vestibular nerve. After complete vestibular compensation, inferior olivary lesion leading to climbing fibre death produced an immediate reversal to the abnormal pre-compensated motor behaviour, implying that the learning had not occurred at the cerebellar cortex. If this were to be the case, the parallel fibre–Purkinje cell synapse, having been modified by climbing fibre activity during the compensation period, should have carried on the learned motor behaviour and vestibular compensation should have been unaltered.

In more recent experiments in cats, a similar phenomenon — loss of a learned motor behaviour after an inferior olivary lesion — was confirmed[7]. Reversible blockage of the olivo-cerebellar system by local anaesthetics rapidly abolished a previously learned compensatory eye movement adaptation to reversing prisms. Because inferior olive blockage abolished this learned adaptation without any delay, and because the learned adaptation reappeared immediately after the cessation of the local anaesthesia, it was concluded that no memory engram was present in the cerebellar cortex; rather, the climbing fibre system seems to be a true afferent system, not a modifier of synaptic efficiency at the Purkinje cell. This suggests that the actual modifiable system is not located in the cerebellum, and so the cerebellum must be considered an important centre for the expression, but not the locus, of the actual motor learning.

There are other general models of cerebellar function, some quite similar to Marr's[8] and others rather different[9,10], of which the differences, virtues and flaws are still being debated. The cerebellum continues to be one of the most dynamic areas of research in neurobiology; it continues to give us new and important clues regarding the function of single cells and the properties of neuronal networks. Furthermore, as we begin to understand the nature of motor coordination — the ability of the brain to transform desired movements (intended movements) into actual movements (motor executions) — the clearer will become the role of the cerebellum in the gamut of brain transactions, and this is the reason for hope! Because some of the models being discussed present formal alternatives to how the cerebellum coordinates movements and because of the theoretical interest which such understanding has in relation to the increasing interest in artificial intelligence and robotics, it is very likely that the cerebellum will continue to offer one of our best chances for obtaining a global understanding of any region of the nervous system. Whatever the outcomes of these efforts, the fact ultimately remains that individuals like Marr appear rather seldom and in losing him we lose more than a good friend and a respected colleague. □

1. Marr, D. *J. Physiol., Lond.* **202**, 437 (1969).
2. Braitenberg, V. & Atwood, R. P. *J. comp. Neurol.* **109**, 1 (1958).
3. Brindley, G. S. *Int. Brain Res. Org. Bull.* **3**, 80 (1964).
4. Ito, M. in *Control of Gaze by Brain Stem Neurons* (eds Baker, R. & Berthoz, A.) 391 (Elsevier, Amsterdam, 1977).
5. Gilbert, P. F. C. & Thach, W. T. *Brain Res.* **128**, 309 (1977).
6. Llinás, R., Walton, K., Hillman, D. E. & Sotelo, C. *Science* **190**, 1230 (1975).
7. Demer, J. L. & Robinson, D. A. *Invest. Ophthalmol. vis. Sci., Suppl.* 20 (1981).
8. Albus, J. S. *Math. Biosci.* **10**, 25 (1971).
9. Pellionisz, A. & Llinás, R. *Neuroscience* **4**, 323 (1979).
10. Pellionisz, A. & Llinás, R. *Neuroscience* **5**, 1125 (1980).

Grand bang

from Paul Davies

WE are all familiar with three states of matter — solid, liquid and gas. Each phase can be reached by heating the previous phase. This is not the whole story, however, as physicists have discovered new phases of matter that arise when temperatures are raised still further. At several thousand degrees a gas will ionize into a plasma and at several million degrees the nuclei fragment into individual protons and neutrons, while at still higher, as yet uncertain, temperatures, the identities of even these nuclear particles are lost. The behaviour of matter is then determined by the quarks that compose the neutrons, protons and other subnuclear particles.

At each successive phase transition, some structure is lost. A melted solid loses its crystalline structure, a boiling liquid gives way to a chaotic mêlée of molecules, and with higher temperatures the molecules, and eventually the atoms themselves break up. All distinction between atoms disappears when the nuclei fragment until finally, even names such as 'neutron', 'proton' and 'meson' become meaningless.

Now some physicists are speculating about yet another phase of matter, beyond the quarks, which is achieved at the amazingly high temperature of 10^{30}K. Under these extreme conditions, matter is smashed into its ultimate building blocks, and theory predicts that even the quarks and leptons lose their individual identities. The fundamental forces that control the structure of the microworld merge into a single unified force, and physics is reduced

---

Paul Davies is in the School of Physics at Newcastle University.

0028-0836/81/220280-02$01.00

Dominance hierarchies among worker ants

from Robert M. May

THE pecking order of barnyard fowls and other birds, and the dominance hierarchies of hamadryas baboons and other primates (including academics), are widely known. One tends to think of such phenomena, however, as being found only among vertebrate species.

In a remarkable recent study, Cole (*Science* **212**; 83, 1981) has demonstrated the existence of linear dominance hierarchies among the workers of the ant species, *Leptothorax allardycei*. These dominance hierarchies, moreover, have all the classic characteristics of their vertebrate analogues: routine displays of dominance, avoidance behaviour and even fighting.

Cole collected colonies of these neotropical myrmicine ants from their habitats among dead sawgrass stems in pine and palmetto scrub (in the Florida Keys) and housed them for observation in glass nests. Encounters between worker ants were then watched and the results tabulated in the conventional matrix form; these interaction matrices comprise totals of 200–400 interactions, over 10–20 hours of observation. In Cole's words, the agonistic encounters begin "with the dominant ant touching the submissive ant with its antennae. There is brief pause before the dominant either rapidly turns or lunges toward the subordinate. It then pummels the gaster, thorax, or head of the subordinate with its mandibles. The encounter may continue, with the dominant ant advancing onto the top of the subordinate, sometimes climbing clear of the nest floor and standing on the subordinate. The typical response of the subordinate is to crouch and freeze, with the antennae drawn back to the sides of the head. It stays in this position until the dominant moves away." The dominance hierarchies thus established are almost perfectly linear, with less than 0.5 per cent of the interactions being 'reversals'. The highest ranking worker accounted for about 43 per cent of all dominance displays, and the second ranking for about 28 per cent; this is typical also of vertebrate dominance hierarchies (see, for example, Wilson *Sociobiology*, Harvard University Press, 1975), and is plausibly explained on the grounds that the higher the rank, the more worthwhile it is to expend effort in defending it.

Cole demonstrates two main advantages that accrue to the dominant workers. First, they are favoured in the exchange of liquid food. Although it is common for liquid food to be exchanged among workers in colonies of social insects, in *L. allardycei* this food transfer is always in one direction. High-ranking workers receive food from low-ranking ones, but do not reciprocate; Cole estimates the probability that chance alone could account for such apparently unidirectional transfer of food to be less than 0.001.

Second, Cole showed that workers laid roughly 20 per cent of the colony's eggs. He did this by removing the queen for 3 days, and feeding the workers food in which Sudan black was mixed; Sudan black is fat soluble and not transferred in liquid food exchange, so that the queen is not contaminated with the stain when she is returned to the colony. As a result, the queen's eggs were the normal white colour, while those produced by workers (amounting, on average, to 21.8 per cent of the total) were stained blue. The haplodiploid sexual system of these social insects means that the workers' eggs (which are unfertilized) produce males; the fertilized queen can produce males or females. Cole further demonstrated that the dominant workers had substantially greater ovarian development, and thus presumably accounted for most of the worker-produced males. In short, *L. allardycei* has effectively two reproductive castes: the queen, who produces all the females, and the high-ranking workers, who produce at least a significant fraction of the males.

The dominance hierarchies in *L. allardycei* constitute a most engaging curiosity. They also hold larger evolutionary implications. As Cole emphasizes: "this social organization . . . has characteristics intermediate between those of the supposed nonsocial ancestors or ants and those of advanced eusocial species. Consideration of the social organization of this species must alter the conception of the reproductive options available to the worker caste of social insects."

Robert M. May is Class of 1877 Professor of Zoology at Princeton University.

---

to its most basic simplicity. Or so the theorists claim.

Temperatures of 10^{30}K are, to say the least, inaccessible to present technology. There was one occasion, however, when matter could have been as hot as this: during the big bang. The creation of the Universe generated enormous quantities of heat, and it is generally supposed that the temperature of the primeval furnace plummetted from unlimited values at the outset to a mere 10^{10}K within one second. For about 10^{-35} seconds (which is brief even by physicists' standards), the temperature could have been high enough for the predicted ultra-simple phase of matter to have fleetingly existed.

Matter, and the forces that control it, are pictured as having frozen out of the featureless primeval soup in a succession of phase transactions, each with progressively more structure. The distinctive differences between leptons and quarks, and between electromagnetic, weak and strong nuclear forces, are regarded as a purely low-temperature differentiation.

When water freezes to ice, it releases a great deal of heat — so-called latent heat — which is why water can seem reluctant to solidify at 0°C. Similarly, in any phase transition, heat is either absorbed or expelled to help drive the transition itself, even without any change of temperature. Calculations suggest that when the quark–lepton phase froze out of the primeval furnace, a large quantity of heat was also released.

This most exotic phase of matter has been studied by A.D. Linde of the Lebedev Physical Institute in Moscow. In a recent paper (*Phys. Lett.* **99B**; 391, 1981), Linde argues that the quark–lepton freeze-out was an event of shattering proportions. What elevates the primeval latent heat release to the status of a 'bang' is the phenomenon of supercooling. It is well known that a very pure vapour can be carefully cooled to below its condensation temperature without liquid drops appearing. A slight disturbance will then lead to a catastrophic appearance of droplets. This principle was used in the so-called cloud chambers to detect the passage of subatomic particles.

Similarly in the primeval Universe, the quark–lepton phase transition did not occur as soon as the temperature fell to the 'condensation' temperature. Because it depends on certain quantum 'tunnelling' processes that require a finite duration, the transition was inhibited until about 10^{-35}s, by which time the temperature was down to a modest 10^{28}K. When the transition finally occurred, in the words of Linde, it took on the features of a great explosion "shaking the whole Universe and changing abruptly its temperature".

The catastrophic release of energy suddenly reheated the cooling Universe and drastically altered the details of its evolution at a most sensitive epoch. Current theories suggest that it was at this stage that the Universe acquired its predominance of matter over antimatter, and that various enigmatic processes occurred which suppressed the production of magnetic monopoles. Linde calls this phase transition 'the grand bang' in honour of the so-called grand unified theories of fundamental forces that predict such a transition. No doubt its reverberations will disturb some of the tidy theories currently being formulated concerning the very early Universe.

0028-0836/81/220281-01$01.00

Malaria vaccine protects offspring

from F.E.G. Cox

THERE can be no doubt that an antimalarial vaccine is one of the major requirements of tropical medicine and a number of laboratories are attempting to develop one. There is equally no doubt that vaccination against malaria is possible and has been achieved using a wide range of experimental model systems, but the protection produced varies in efficacy and all the methods so far described have serious drawbacks.

In the life cycle of the malaria parasite in the human host, the infective stages injected by the mosquito are the sporozoites, which circulate in the body for about half an hour before entering the liver where a massive phase of multiplication occurs. This results in the production of thousands of merozoites which invade red blood cells where another multiplication phase, erythrocytic schizogony, takes place. The products of this phase are further merozoites which invade fresh red blood cells in which erythrocytic schizogony again occurs and this process is repeated until the host dies or recovers. Eventually sexual stages are produced and these are taken up by and infect the mosquito vector[1].

The various stages in the life cycle differ antigenically from one another; thus as the immunity produced is stage specific, a blockbuster approach to immunization is not feasible. At present the main targets for vaccines are the sporozoites, the merozoites, the erythrocytic stages and the gametocytes[2]. The use of irradiated sporozoites as a vaccine has been largely confined to rodent models but with considerable success. Adult mice can be protected by the intravenous injection of irradiated sporozoites[3]. The protective antigen is a surface antigen that can be demonstrated with the use of monoclonal antibodies[4] and monoclonal antibodies against this antigen protect mice against challenge with the homologous species[5]. On page 331 of this issue of *Nature* Orjih and colleagues report that these findings have been extended to young mice and the offspring of immunized mothers.

Briefly, adult mice given one to four immunizing doses of irradiated sporozoites intramuscularly were poorly protected against challenge with live sporozoites several weeks later. Young mice 2–14 days old, on the other hand, were well protected. The fact that adult mice can be protected by the intravenous route but not the intramuscular one is interesting but what these experiments do show is that young animals can be immunized in an acceptable way. Of even greater importance is the fact that when adult female mice were immunized intravenously with irradiated sporozoites, over half the offspring were protected against challenge and similar levels of protection were found in mice born to normal mothers but fostered onto immunized ones.

These results are particularly important because, in many parts of the world, malaria is essentially a disease of children. In endemic areas, babies acquire immunity passively from their mothers and are then protected during the critical first few months of life. Control schemes, however, upset this balance and it is in such situations that a vaccine that protects young children is likely to be of most use. Another important aspect of these results is that similar findings may apply to other vaccines such as the merozoite one being developed at Guy's Hospital in London[2], and recently described improvements in the *in vitro* cultivation of human malaria parasites[6,7] and adjuvants[8] have moved such a vaccine a further step forwards. The merozoite vaccine is known to protect monkeys but for obvious reasons the possibility of transfer of immunity to their offspring has not yet been investigated.

Those who have advanced the cause of a malaria vaccine have never suggested that a totally effective vaccine would ever be possible but have argued that a vaccine could produce better protection than a natural infection and could ameliorate the course of the disease. These experiments with ***Plasmodium berghei*** appear to vindicate this cautious approach. However, before too much euphoria sets in it must be remembered that the passage of immunity from mothers to offspring differs markedly from species to species and what happens in a mouse or monkey need not necessarily happen in man. □

F.E.G. Cox is Professor of Zoology at King's College London.

1. Cox, F.E.G. *Biologist* **28**, 9 (1981).
2. Cohen, S. *Proc. R. Soc.* **B203**, 323 (1979).
3. Orjih, A.U. & Nussenzweig, R.S. *Am. J. trop. Med. Hyg.* **29**, 343 (1980).
4. Yoshida, N. *et al. Science* **207**, 71 (1980).
5. Potocnjak, P. *et al. J. exp. Med.* **51**, 1504 (1980).
6. Butcher, G.A. *Ann. trop. Med. Parasit.* **75**, 111 (1981).
7. Scott-Finnigan, T.J., Wilson, R.J.M. & Williamson, J. *Trans. R. Soc. trop. Med. Hyg.* **75**, 292 (1981).
8. Siddiqui, W.A. *et al. Nature* **289**, 64 (1981).

Continental swells

from Peter J. Smith

As newly formed lithosphere spreads away from oceanic ridges it cools and contracts, leading to a subsidence of the ocean floor that increases generally with age. Indeed, as Sclater and colleagues have shown (*J. geophys. Res.* **76**; 7888, 1971), there is a simple relationship between lithospheric age and water depth, which may be regarded as the first-order description of ocean floor elevation. However, as with most geological phenomena, there are also second-order and even third-order effects — local deviations from the general rule. Chief among these are the broad topographical swells, typically 1,000 km or so across and about 1 km high, associated with zones of extensive intraplate volcanism such as the Canaries, the Cape Verde Islands and Hawaii, as well as the topographical highs representing aseismic ridges such as the Rio Grande Rise and volcanically long extinct regions such as Bermuda.

Cochran and Talwani (*Geophys. J.* **50**; 495, 1977) have recently shown that both the volcanically active topographical features (known in certain quarters as hot spots) and their inactive counterparts are associated with positive gravity anomalies of intermediate wavelength — an observation that poses certain problems of interpretation, especially if volcanic activity alone is to be regarded as the lithospheric manifestation of mantle plumes. Similar gravity effects suggest common underlying causes; and yet it is not entirely clear how topographical anomalies associated with volcanic activity, whether plume-related or not, could be supported by the same mechanism as, say, aseismic ridges. Detrick and Crough (*J. geophys. Res.* **83**; 1236, 1978) may or may not be right, therefore, in attributing the 'active' topographical swells to reheating of the lithosphere.

Whatever the mechanism may be, what produces little more than a topographical swell in oceanic lithosphere may produce rather more complex results in continental lithosphere. This supposition has led Crough (*Geology* **9**; 1, 1981) to examine the consequences of a possible topographical swell moving across parts of the United

---

Peter J. Smith is in the Department of Earth Sciences, The Open University, UK.

0028-0836/81/220282-02$01.00

States and Canada. From the known drift of North America and information about dated igneous intrusions, it is clear that from at least 160 to 100 Myr ago the continent passed northwestwards over the hot spot now near Great Meteor Seamount in the eastern Atlantic. If that hot spot, by whatever means, produced a topographical swell throughout its journey, *a priori* it should have had an influence on continental evolution in terms of elevation, thermal history and patterns of erosion along a track from at least 50°N latitude in Canada southeastwards down to the New England coast.

Crough's case for suggesting that such an influence was indeed exerted begins with the observation that the hot spot trace runs approximately along the northern limit of the Palaeozoic sedimentary cover in eastern North America. To the south of the track lies the Appalachian basin; to the north lies the largely exposed Canadian Shield. There are numerous lines of evidence, however, to suggest that the sediments of the Appalachian basin once extended much further to the north through southern Canada.

First, there is the circumstantial evidence provided by the nature of the existing basin sediments of late Ordovician and younger age, foreland clastic deposits derived from Palaeozoic mountains to the east. It is reasonable, claims Crough, to suppose that these mountains would also have given rise to enough sediment to bury the shield to the west. Second, the present outliers of Palaeozoic (Ordovician–Devonian) sediment in the shield area imply a once more extensive cover. Third, isopach maps of the remaining sediment in the Appalachian basin indicate no thinning towards the north, as might have been expected had the present northern edge of the basin always been the limit of sedimentary cover.

Fourth, lithofacies relations show the present northern margin of the basin to be defined by post-depositional erosion. There is no evidence of a clastic source along that margin, for the clastic ratio of the basin sediments increases from west to east with little north–south variation; and in any case, palaeocurrent indicators generally demonstrate sediment transport towards the north-west and west rather than from the north. Finally, conodont geothermometry establishes that those few sediments remaining in southern Canada and northern New York were once deeply buried, perhaps beneath up to 4 km of overburden.

So what happened to this once more extensive sediment cover to the north? Obviously it was eroded away; and it is Crough's contention that the erosion occurred as the hot spot-induced uplift, at a maximum (and hence producing maximum erosion) along the axis of the hot spot track, moved down from Canada towards New England. In other words, a hot spot moving relative to an overlying continent can have a far more profound effect than simply generating volcanic activity at the surface — although there is that too, as the kimberlites and alkalic complexes, not to mention the offshore New England seamounts, testify.

The chief information for the timing of the passing uplift and erosion comes from apatite fission-track ages of the basement along the supposed hot spot track. Apatite acts as a maximum thermometer, recording the time at which the mineral last cooled below 100°C, which, if Crough is right, would be the time at which the removal of the Palaeozoic cover allowed cooling from burial temperatures. The many apatite fission-track ages already available from the region are almost all Cretaceous to early Tertiary and thus consistent with the passage of the Great Meteor hot spot.

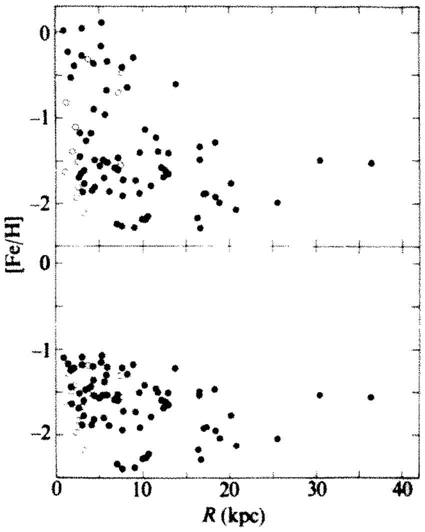
Crough's interpretation differs from some others; Dewey and Kidd (*Geology* **2**; 543, 1974), for example, attributed variations along the Appalachians to processes occurring at the time of orogeny. Insofar as Crough's study applies to a particular region, it must be judged by others familiar with the regional geology. But if hot spot-induced continental uplift proves to be common (and only examination of the loci of other subcontinental hot spots will decide this issue), it will presumably force a general re-evaluation of the style, and perhaps timing, of much continental erosion. □

A revision of the globular cluster metallicity scale

from Barry F. Madore

FEW probes of the spatial, chemical and dynamical evolution of our Galaxy have held such a pre-eminent position as have the globular clusters. Being the surviving subunits of the early collapse of our Galaxy, these bright, self-contained systems of old stars have long been used by astronomers to calibrate the size and composition structure of our own and other galactic systems. However, this legendary structure is coming under rapid reappraisal with mounting evidence that the basic calibrations of the metal content of globular cluster stars may be in error.

The collapse of the calibration is a direct result of advances in instrumentation and the observational efforts of Catherine Pilachowski, George Wallerstein, Ron Canterna and others from the University of Washington, as well as from the studies of Judith Cohen at the California Institute of Technology. Hints that the metallicity of the most metal-rich globular clusters in our Galaxy had been overestimated first surfaced at the Advanced Study Institute on Globular Clusters*, and these results were later presented in a refined form (C. Pilachowski, C. Sneden & R. Canterna, p.467; J. Cohen, p.385, *Star Clusters*, ed. J. Hesser, Reidel, Holland, 1980). The most metal-rich globular clusters are now cited as being some five times more metal poor than earlier calibrations had suggested. The calibration for the metal-poor globular clusters has remained relatively unaltered.



The logarithmic metallicities of globular clusters as a function of galactocentric distance, *R*. The upper panel shows the metallicity gradient obtained when using the old metallicity calibration while the lower panel shows the flattening of this distribution when the newest calibration is applied. Open circles are for globular clusters with uncertain distances. (Taken from *Astrophys. J.* **241**; 606, 1980.)

Robert Zinn has reviewed the implications of this revision for models of early galactic chemical enrichment (*Astrophys. J.* **241**; 602, 1980). The oft-cited 'metallicity gradient' for globular

---

*The Advanced Study Institute on Globular Clusters was sponsored by NATO at Cambridge, England in August 1978.

Barry F. Madore is at the David Dunlap Observatory, Ontario, Canada.

0028-0836/81/220283-02$01.00

clusters as a function of radial distance from the centre of our Galaxy is shown in the upper panel of the figure; this is using the old metallicity scale. The gradient is clearer in the upper envelope of this distribution. At face value this plot indicates that the most metal-rich globular clusters preferentially inhabit — and by implication, were preferentially formed in — the inner regions of our Galaxy. Metal-poor systems formed in a more extended (earlier) configuration. These data supported the canonical collapse theory of a well mixed galactic enrichment model. However, as illustrated in the lower panel of the figure, the situation is significantly modified by adopting the new metallicity scale: the gradient is practically eliminated and with it the simple theories of galactic collapse.

Even without the gradient, one is still left with a significant spread in metallicity at any radial distance and even among globular clusters of the same metallicity there is the well known problem of 'the second parameter phenomenon'. This phenomenon is manifest by the fact that although some aspects of the distribution of globular cluster stars in the luminosity–temperature plane (such as the height of the cool red giant branch) are a strong function of metallicity, other features (such as the relative population of stars on the red and blue side of the horizontal branch) are independently a function of some as yet unspecified parameter.

Zinn has attempted to provide a unifying solution to these problems using a modified version of a model of galaxy evolution recently proposed by Leonard Searle and himself (*Astrophys J.* **225**; 357, 1978). In a model of cluster formation and destruction, which encompasses the build-up of the galactic disk, the formation of satellite galaxies to our own Milky Way and the deposition of a retinue of surviving globular clusters, Zinn suggests that the second parameter might be age. If so, the clusters in the outer parts of our Galaxy would be some three billion years younger than those in the central regions — a situation almost completely opposite to earlier thoughts on the matter.

The issue is certainly not closed and more models, both of individual globular cluster evolution and the formation of the galactic system, are certain to be forthcoming. Indeed, a recent theoretical paper by Demarque and McClure (*Astrophys. J. Lett.* **242**; L5, 1980) concludes, "through a process of elimination", that helium abundance, not age, is a dominant second variable in globular cluster evolution. While high helium abundance would replace older ages for the globular clusters, with this interpretation would come the need for a source of helium abundance variations. A generation of supermassive stars could provide this extra helium at various places in the galaxy but the ultimate reasons for the systematics of their occurrence have yet to be speculated on. ☐

Cortical effects of monocular deprivation: suppression or deafferentation?

from J. Anthony Movshon

It is eighteen years since Wiesel and Hubel[1] showed that depriving a kitten of vision in one eye causes that eye to lose virtually all effectiveness in activating visual cortical neurones. The developmental mechanism underlying this shift in cortical eye dominance has long been known to depend on a competitive interaction between the eyes[2], but the neurophysiological basis of the deprived eye's inability to influence cortical neurones remains a matter for debate. On the one hand, there is abundant evidence that monocular deprivation causes a radical reduction in the extent of cortical terminations of afferents originating in the deprived eye[3]. On the other hand, there is reason to suppose that residual inputs from the deprived eye remain, but are functionally suppressed by inhibitory influences from the normal eye[4].

A mechanism that could be responsible for this suppression is intracortical inhibition mediated by synapses using γ-aminobutyric acid (GABA) as a neurotransmitter[5]. A few years ago, Duffy and his colleagues[6] claimed that intravenously administered bicuculline (a GABA antagonist) could reveal an important input from the deprived eye in monocularly deprived cats. They proposed that a GABA-mediated inhibition contributed significantly to the cortical effects of monocular deprivation. There were, however, a number of important defects in these experiments — most obviously, intravenously administered bicuculline may act at many sites other than the cortical one in question. It is thus welcome that Sillito and his colleagues (*Nature*, this issue, p.318) have now repeated Duffy's experiment using the much more sensitive method of microiontophoresis to deliver minute quantities of bicuculline to the immediate vicinity of the recording electrode. Sillito and co-workers also control for the possibility that it is not the block of GABA *per se*, but rather the overall increase in excitability that results from it, that might unmask the deprived eye's input. Their results show that in a substantial minority of neurones recorded from monocularly deprived cats, bicuculline reveals a GABA-suppressed input from the deprived eye; this input rarely matches and never exceeds in effectiveness the input from the normal eye.

Although Sillito's results seem to support a suppression model of the cortical effects of deprivation, a separate study from his laboratory[7] makes it unlikely that this suppression is specific to deprived cats: apparently even in normal cortex, about half the monocularly activated neurones change eye dominance during bicuculline administration. It would seem that whatever mechanisms mediate this inhibition in normal cats continue to function in deprived cats, and that the inhibition is not a consequence of the deprivation.

There remain some important issues addressed neither by Duffy's group nor by Sillito's. The most important of these concerns the laminar location within the cortex of neurones that show suppression effects. The anatomical data concerning the redistribution of afferents from the two eyes after deprivation apply only to layer IV, the layer within which the bulk of the specific afferents to cortex terminate. In normal animals, the eye dominance of a 'column' of cortical tissue — extending both above and below layer IV — is determined largely by the eye dominance within the adjacent patch of layer IV. In deprived animals, however, eye dominance outside layer IV may depend on other factors; deprivation of late onset can cause marked changes in cortical dominance outside layer IV without altering the distribution of inputs to layer IV[8]. It thus appears that two mechanisms of deprivation must exist, one that determines the distribution of afferents within layer IV, and a second that determines the relative effectiveness of the output from different patches of layer IV to other cortical layers. The extra-layer IV mechanism might be, in part, a suppressive one, and it is thus important to know the laminar distribution of neurones showing GABA-mediated inhibition from the normal eye. Moreover, a difficulty arises if this suppressive interaction is confined to particular layers: interocular suppression effected in one layer might have a

---

J. Anthony Movshon is Associate Professor of Psychology, New York University.

1. Wiesel & Hubel *J. Neurophysiol.* **26**, 1003 (1963).
2. Wiesel & Hubel *J. Neurophysiol.* **28**, 1029 (1965).
3. Hubel *et al. Phil. Trans. R. Soc.* **B278**, 377 (1977); Shatz & Stryker *J. Physiol., Lond.* **281**, 267 (1978).
4. Kratz *et al. J. Neurophysiol.* **39**, 501 (1976).
5. Iversen *et al. J. Physiol., Lond.* **212**, 519 (1971).
6. Duffy *et al. Nature* **260**, 256 (1976).
7. Sillito *et al. Expl Brain Res.* **41**, 1 (1980).
8. Hubel *et al. Phil. Trans. R. Soc.* **B278**, 377 (1977).

0028-0836/81/220284-02$01.00

powerful influence on neurones — recorded in other layers — that receive input from the 'suppressed' layer. Yet microiontophoretically applied bicuculline would not spread far enough to be effective in layers remote from the recording site. Thus Sillito and colleagues' estimate of the proportion of neurones influenced (directly or indirectly) by GABA-mediated inhibition might well prove too low.

In any case, it is evident that both the redistribution of afferents within layer IV and a degree of interocular suppression (perhaps outside layer IV) contribute to the cortical effects of monocular deprivation. While the available evidence suggests that interocular suppression plays a secondary role, resolution of the issue must await detailed study of the laminar location and intracortical connectivity of 'suppressed' neurones, as well as attention to potential inhibitory neurotransmitters other than GABA. Plainly, Wiesel and Hubel's straightforward observations of eighteen years ago opened the door on complex features of cortical organization. □

Global warming?

from T.M.L. Wigley, P.D. Jones and P.M. Kelly

In 1975, Lamb and his colleagues[1] noted that the downward trend of Northern Hemisphere surface temperatures, which had begun in the early 1940s, had come to an end — at least in the northern North Atlantic. Since then, the direction of the global temperature change has been a controversial issue. Some argue, justifiably, that we cannot be sure of global-mean conditions because there are such large gaps in the data coverage, particularly over the ocean and in the Southern Hemisphere. However, it can be argued that, because temperatures are more uniform over the oceans than the continents, large oceanic data gaps can be tolerated (see Barnett[2] for an opposing view). Further controversy lies in whether or not surface temperatures are meaningful at all. Perhaps a better measure of large-scale fluctuations in atmospheric temperature is the column-mean temperature, as determined, for instance, by the 1,000 to 500 mbar height difference. Doubts have been expressed by Parker[3] as to the reliability of these data but, provided the limitations are taken into account, considerable insight into global temperature changes can be gained.

Much useful information can be gained from recent Russian publications. Borzenkova and his colleagues[4] gave a series of annual mean temperature values averaged over the latitude belt 17.5–87.5 °N for the period 1881–1975 which show a marked warming since 1964. Despite criticisms[5], Vinnikov[6] has now convincingly shown that the data are homogeneous and do not suffer from discontinuities apparent in earlier work. Vinnikov also pointed out that the warming trend they had detected was certainly not spatially 'universal' . . . "variations of air temperature at the surface of the Earth may have different signs over extensive areas of the globe in specific seasons of the year, even against the background of a significant rise in average global characteristics" (p. 97). Thus, while an overall warming may have occurred, quite large regions may still exhibit a cooling trend. Although the trend revealed by Borzenkova appears, therefore, to be well established, one comment[5] must be accepted — that we do not know whether the warming is liable to continue into the future. All of this, of course, begs the questions of what is meant by 'trend', and how one determines the beginning of a trend, given the large year-to-year variability.

What has happened since 1975? A recent Russian paper[7] gives results up to and including 1978, and we[8] have independently made estimates of Northern Hemisphere average surface temperatures up to February 1981. Our own results, based on very much the same data as used by the Russians, naturally correlate highly with theirs. The data show 1972, 1974 and 1976 to be quite cold years, agreeing with the estimates of Barnett[2] and Angell and Korshover[9]. In fact, the data show 1972 to have been the coldest year since 1917. Vinnikov *et al.*'s data agree for 1972 and 1976, but they show 1974 as a warm year. A cooler 1974 and a cool year in 1976 reduce the magnitude of Borzenkova *et al.*'s 1964–1975 warming trend. However, 1977 was a warm year (the warmest since 1953) and both 1979 and 1980 were also relatively warm. The warming trend appears, therefore, to be continuing; but conditions are still considerably cooler than the balmy period between the late 1930s and the early 1950s.

Perhaps the most outstanding feature of recent years has been the 1981 Northern Hemisphere winter (December 1980–February 1981). This was the warmest winter on record (records began in 1881), equalling the value for 1944. The next three warmest winters are, in order, 1926, 1958 and 1980. In the past two years we have, therefore, experienced two remarkably warm winters, contrasting sharply with the winters of 1969 and 1972 which were the two coldest winters since 1917. For the 1981 winter, especially in January and February, the maps published by the Free University of Berlin (*Berliner Wetterkarte*) show remarkable warm anomalies over northern Canada and Alaska, and over north-central Siberia — with a much smaller-magnitude cool anomaly over the North Atlantic. Many places in northern Canada and Alaska recorded record warmth in January, some places being more than 16°C above the 1931–60 'normal'.

The warmth of the past two winters may be reflected in Northern Hemisphere snow cover. In a recent paper in *Nature,* Matson and Weisnet[10] described changes in snow cover, based on satellite measurements, over the period 1967–1980. Over North America, a rising trend up to 16.9×10^6 km^{-2} in 1979 was followed by a sharp drop in 1980 (15.1×10^6 km^{-2}). The 1981 value is even lower, 13.9×10^6 km^{-2}, the lowest on record. For Eurasia the 1980 and 1981 values are both low (but not as low as the 1970 value), while for the whole hemisphere the values for 1980 and 1981 are 38.9×10^6 km^{-2} and 38.0×10^6 km^{-2}, values surpassed only by the 1970 value (37.2×10^6 km^{-2}) (1981 data from Michael Matson, NOAA/NESS, Washington, DC 20233). In spite of the correspondence between recent winter warmth and low snow cover, the overall correlation between these two variables is low. For the first 13 years of record the correlation coefficient between hemispheric winter snow cover and temperature was −0.03. When 1980 and 1981 are included this rises to −0.40!

It is difficult to know the significance of these recent events, but it is certainly premature to associate them with the effects of increasing atmospheric carbon dioxide. At present, the high inter-annual variability of surface temperatures (that is, the 'noise' of natural climatic variability) makes it impossible to identify the apparent trend with any particular causal mechanism. We note also that a similar warming trend has occurred before, between the late 1910s and around 1940. However, a few more consecutive warm winters, with continuing warmth in the other seasons, will doubtless give climatologists pause for thought. □

T.M.L. Wigley, P.D. Jones and P.M. Kelly are in the Climatic Research Unit, School of Environmental Sciences, University of East Anglia.

1. Lamb, Malmberg & Colebrook *Nature* **256**, 479 (1975).
2. Barnett *Mon. Weath. Rev.* **106**, 1353 (1978).
3. Parker *Met. Mag.* **109**, 129 (1980).
4. Borzenkova, Vinnikov, Spirina & Stekhnovskii *Meteorologiya i Gidrologiya* **1976**(7), 27 (1976).
5. Rubinshtein *Meteorologiya i Gidrologiya* **1977**(9), 106 (1977).
6. Vinnikov *Meteorologiya i Gidrologiya* **1977**(9), 110 (1977).
7. Vinnikov *et al. Meteorologiya i Gidrologiya* **1980**(6), 5 (1980).
8. Jones & Wigley *Climate Monitor* **9**, 43, 86, 109, 147 (1980).
9. Angell & Korshover *Mon. Weath. Rev.* **105**, 375 (1977).
10. Matson & Weisnet *Nature* **289**, 451 (1981).

0028-0836/81/220285-01$01.00

Why mammalian gametes don't mix

from J. Michael Bedford

THROUGHOUT the Metazoa the reproductive isolation of a species is consistently reflected in the restricted affinities of sperm and egg. The possibilities for interspecific fertilization and the precise mechanisms that restrict this among mammals, including man, have been less than clear, but an understanding of the basis of this restriction is gradually emerging. An ability to dissect or overcome the specific barriers to cross fertilization could help us come nearer to a precise analysis of the mechanisms of the normal event. More practically, it may also soon offer an avenue for clinical evaluation of the fertilizing ability and even the chromosomes of human sperm. Considering the sequential steps of the conception process allows us to summarize what we know of the factors that prevent easy union of sperm and eggs from different mammals, and why further study of this should be rewarding.

The exact way that a specificity of sperm and eggs is achieved may vary. Sperm of teleost fish, for instance, must negotiate the specialized environment created around a small hole in the egg shell. Usually, however, organisms practicing external fertilization seem to depend primarily on elements in the egg coat and sperm surface whose 'fit' dictates a specificity of penetration; one that for sympatric marine species prevents the sea becoming a gamete mixing pot. Since the behavioural preliminaries to ejaculation and insemination could preserve a specificity of fertilization in mammals, an extreme individuality of the character of their gametes might seem unnecessary. Nonetheless, despite changes in the character of both gametes in evolutionary transition from external fertilization to the rather complex interaction that characterizes fertilization in placental mammals, such an individuality persists. Cross-species fertilization can occasionally follow artificial insemination, or the mating that sometimes occurs in confined situations between related mammals. Further development then depends on the harmony of gene expression in the embryo and on its intrauterine relationship with the mother — one favoured possibly by the superficial rather than the intimate type of placenta. Such embryos often die during pregnancy (ferret x mink; goat x sheep; rabbit x hare), but others survive to adulthood as healthy individuals that may be wholly sterile (horse x zebra) or not (ox x bison). In general, however, mammalian gametes tend to express their specific character at several points in the process of conception. As a consequence, fertilization will not occur readily even between members of the same Family such as the rat and mouse, let alone among mammals that differ more widely.

Preliminaries in the female. Incompatability could first be expressed in the foreign tract by failure of sperm to be transported to the site of fertilization in the Fallopian tube, to survive as viable cells, or to be capacitated (mammalian sperm are not able to fertilize immediately after ejaculation and must undergo a final maturation, capacitation, in the female tract). Little is known of sperm function in the native tract, and so the factors affecting foreign spermatozoa are hardly definable. There is no clear pattern to their longevity: occasionally it can be prolonged beyond that seen in the homologous tract (rabbit spermatozoa in the estrous rat uterus), but foreign spermatozoa more often die before they would in a native situation and before native spermatozoa in the same uterus or oviduct. Nor does the persistence of motility guarantee transport to the site of fertilization. As shown by the behaviour of human spermatozoa inseminated into the rabbit, maintenance of their motility in the vagina is no ready passport through the cervix to the uterine cavity. The Fallopian tube likewise tends to select the native male gamete from a mixed population of motile cells placed in the uterus, and one can only suspect involvement of the specific surface character of the sperm cell at either point of selection in the tract. On the other hand, capacitation does not appear to be an obstacle when motility is preserved and spermatozoa can experience the tubal environment. Moreover, the time required for capacitation seems more a function of the specific identity of the sperm than of the environment.

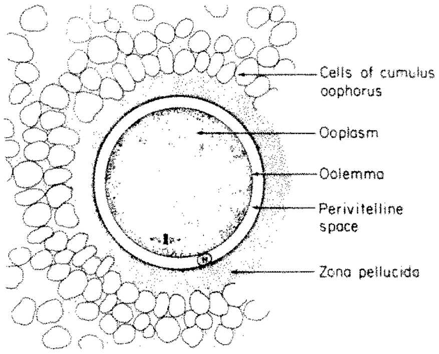


Fig. 1 Diagram of a typical mammalian egg soon after ovulation that illustrates the various vestments through which the sperm must pass before being incorporated into the ooplasm.

Sperm penetration of egg vestments. For gamete union to occur, sperm must first pass through the substantial barriers around the newly ovulated egg (see Fig. 1). The specificity of the immediate preliminaries to the union of mammalian gametes is difficult to elucidate. At the site of fertilization, little is known of the nature of the elements that promote the acrosome reaction, the role and specificity of action of acrosomal enzymes, and the exact function of the specific binding between the reacted sperm head and zona surface. Whether once capacitated, spermatozoa are then triggered to react by a specific signal or do so spontaneously is not resolved. The response of capacitated spermatozoa around foreign and native eggs in the same milieu might illustrate whether the egg coat provides a specific reaction (as it does in sea urchin) and if not, whether the cumulus/corona cell complex expresses a selective resistance towards reacted foreign sperm. That the zona pellucida does so is impressively illustrated by the many successes in interspecific fusion that follow its removal. How it exerts this barrier role still eludes us, again because the precise mechanism of homospecific penetration of the zona is largely unexplained. There is an element of 'Catch-22' in that uncapacitated spermatozoa (which cannot undergo the acrosome reaction) will often attach to the zona pellucida non-specifically in associations incompatible with penetration; but in attaining the capacitated state (in which they undergo the acrosome reaction), sperm probably undergo a surface change which imposes specificity in attachment/binding between the zona substance and sperm plasmalemma. At all events it is probable that specific entities, possibly glycoproteins, are exposed on the surface of the capacitated sperm cell membrane overlying the acrosome, intact or reacted — and that these bind effectively to complementary elements at the surface of the zona pellucida. In the mouse, one of three major glycoproteins that form the zona pellucida is implicated as the receptor through which specific or functional ligands are established. Is the tenacity of this binding the essential key? The answer must await resolution of the means by which spermatozoa normally penetrate the zona substance. The thrashing bound spermatozoon may need the stable union of the reacted peri-acrosomal plasmalemma with the zona surface as a base from which to intrude into the substance of the zona. It does not seem good sense, however, that the now-exposed inner membrane of the acrosome should bond with receptors throughout the zona, since that surface must oscillate and slide freely in cleaving a penetration slit. Neither of the prime factors being considered — a motive force imparted by the

---

J. Michael Bedford is Uris Foundation Professor of Reproductive Biology, Cornell University Medical College.

0028-0836/81/220286-03$01.00

vigorous tail, and lytic action of residual acrosomal enzyme(s) — carries an obvious element of specificity.

Sperm–egg fusion. Examples have been reported (white-foot deermouse spermatozoa with rat eggs) in which spermatozoa penetrating the zona pellucida then fail to fuse with the mature oolemma. However, the general role of the zona rather than the oolemma as a prime controller of interspecific fertilization is evident from the wider interactions that follow zona removal. Fusion in eutherian mammals differs from that in other animals, the fusion site being a restricted region of sperm plasmalemma over the stable equatorial segment of the acrosome and not the inner acrosomal membrane as in invertebrates and the few other vertebrates studied. However, these special features are maintained in interspecific fusion among mammals and, according to species, removal of the zona simply destroys much of the exclusiveness of the oocyte, allowing combinations as unlikely as pig with hamster. This does not imply open house at the oolemma. Although the case in the hamster, it is not so in other species which show a variability without clear rules. Neither are the interspecific affinities necessarily equivalent for both gametes. Zona-less mouse oocytes display little general plasticity and are penetrated rarely by rat sperm, yet the reciprocal cross is achieved with much success. Fusion with a foreign sperm generally leads to 'activation' of the egg and completion of the second maturation division of meiosis. However, activation is not an invariable sequitur of fusion and it may fail in certain combinations (for example, rat or hamster spermatozoa associating with mouse eggs, or hamster spermatozoa with rat eggs), though why is not known.

Nuclear events. Is there a specificity also to sperm chromatin and the events that bring its transformation in the post-fusion phase? The answer is a qualified no. Unlike histones, the basic nuclear proteins of spermatozoa are not highly conserved, differences of composition being evident even within the Eutheria. Nonetheless, sperm nuclei in foreign oocytes tend to decondense and often complete the process that transforms them to a large pronucleus appropriate for union with that of the egg. Then the male pronucleus may closely resemble that of the host; for instance, when formed from the pig sperm head in the hamster oocyte it visibly mimics the pronucleus formed there by the hamster spermatozoon. This is perhaps not surprising since autoradiographical and immunological studies suggest that pronucleus formation involves a substitution of sperm protamine by proteins contributed by the egg. There are exceptions that imply a specificity of male pronucleus development; for example, rabbit sperm heads show little response once inside mouse eggs. Though untested, inhibiting factors that might operate in such instances include an absence of egg activation, failure to disperse the investing nuclear membrane and rostral shroud of acrosome and oolemma, or direct incompatibility of sperm chromatin and ooplasmic decondensation factor(s).

Human genetics. Although a viable hybrid has resulted from the mating of a siamang with a gibbon, two closely related apes, there is little known of the interspecific possibilities among primates other than man. The limits of the receptivity of the human oocyte are essentially untested, but while more is known of human spermatozoa the specificity of their function offers little for the fantasists of science fiction or of zooerasty. The evolution of man (there is little evidence for other hominoids) has been accompanied by restrictive change in the affinities of the sperm surface, limiting to an unusual degree the substrates to which it will adhere, capacitated or not. Whereas non-specific attachment of uncapacitated spermatozoa can occur in many mammals, human spermatozoa seem to display almost no affinity for the foreign zona pellucida, even that of sub-hominoid primates such as baboon (see Fig. 2), rhesus and squirrel monkey. Human spermatozoa will attach avidly, however, to the oocyte of the gibbon, *Hylobates lar*, and when capacitated will penetrate the granulosa vestment and zona pellucida. It might be expected that a similar compatibility exists with and perhaps between other hominoid apes, but the developmental potential of ape oocytes penetrated by human spermatozoa is unknown. Experience from other hybrids involving relationships similarly distant to that of man from any of the apes suggests their early failure might be anticipated.

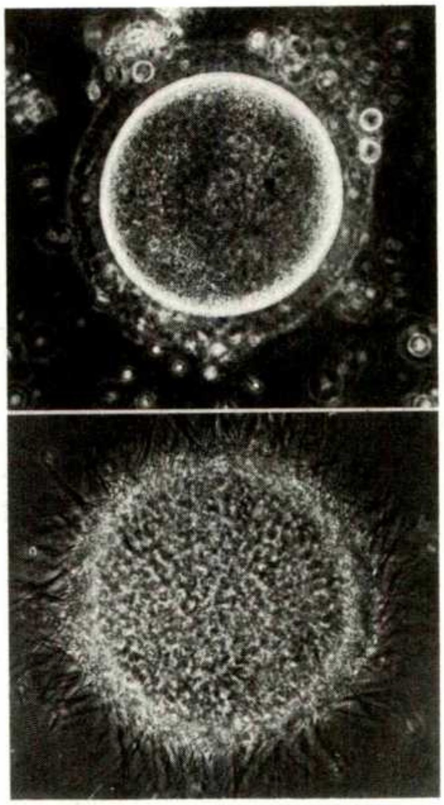


Fig.2 This shows a baboon oocyte (top) and a gibbon oocyte (bottom) that have each been exposed to a similar concentration of human sperm in the same conditions. Capacitated or not, human sperm show no affinity for the zona pellucida of the baboon — only a few sperm are lying adjacent to the zona in a random fashion; by contrast, they adhere avidly to the zona of the gibbon — a true ape — and when capacitated will penetrate the zona.

Clinical uses. Clinicians now seek means of assessing directly the functional quality of the human ejaculate. This would be of great value in diagnosis of the male role in suspected infertility and in evaluating the efficacy of potential male contraceptives that do not deplete the ejaculate of sperm. Although *in vitro* fertilization tests omit a critical element — assessment of the ability of sperm to reach the fertilization site in a viable state — there is great interest in developing this approach to human sperm function. The limited affinities of human sperm mean that the only two options are intact human follicular oocytes that have no capacity for development and/or the zona-less oocyte of the hamster. The first has been shown to be a usable though difficult system in which to test the aspects of human sperm function needed for zona penetration. In the second, an ability of human spermatozoa to undergo an acrosome reaction, fusion and then male pronucleus formation is reflected faithfully in their interaction with the zona-less oocyte of the hamster (though not of mouse or rat).

The human/hamster hybrid system appears a serious possibility for diagnosis of sterility. In some couples with otherwise unexplained infertility, sperm of the male partner fail consistently to enter hamster oocytes after *in vitro* capacitation, even though the suspect sample cannot be distinguished by standard examination from fertile semen. The sensitivity of that system can be increased, moreover, by marking the sperm of the suspect and the normal control samples with different innocuous fluorescent dyes, permitting direct comparison of their fertilizing potential. Unfortunately, while a negative result is of value, the lack of any test of zona penetration makes a positive finding — sperm entry into the zona-less hamster oocyte — an incomplete indicator of sperm normality. The limited evidence supports this last view in that an ability to enter the hamster oocyte and to penetrate the zona of the human follicular oocyte are not always shared; human sperm samples performing in either assay do not always do so in both.

It may prove at least as significant clinically that in human/hamster hybrid fusion the chromatin of the human sperm head can be transformed to a normal male pronucleus and beyond. Culture in the presence of colcemid, an inhibitor of spindle formation, has been shown to allow the preparation of the karyotype of the penetrating spermatozoon (variably more than one per egg), of a clarity sufficient for evaluation of the numbers and the character of the sperm chromosomes. With wider establishment of this technique, studies of the possible incidence with which spermatozoa act as primary agents of the chromosome imbalance so common in early human embryopathies, and the degree to which environmental hazards can affect sperm chromosomes, would seem to be on the horizon. Whatever the conditions of culture, however, the 'humster' zygote fails after one cell division at the two-cell stage. Even in a world where bacteria are coerced to produce human insulin, small boys with furry brown ears are not a foreseeable prospect!

Understanding natural killer cells

from Peter Beverley

NATURAL KILLER (NK) cells were first identified by their ability to kill, without deliberate prior immunization, certain tumour target cells grown *in vitro*. Otherwise they were defined only in negative terms: that is, they were not thymus-derived (T) or bone marrow-derived (B) lymphocytes, nor were they adherent or phagocytic. Though their origins were thus rather mysterious, they were at first widely presumed to be the mediators of immunosurveillance against tumours[1]. Subsequent studies have suggested however that NK cells may be related to either the T[2] or myeloid[3] lineages and may have an immunoregulatory role[4]. Furthermore, several authors have demonstrated heterogeneity of NK effector cells, and that different target cells are differentially susceptible to these effector subpopulations[5]. In the face of this apparent confusion can it now be said that there is such an entity as an NK cell and is it possible to define its biological role?

Examination of mutant animals deficient in various cell types, such as the nude or beige mouse[2], or of humans with genetically determined diseases[6], has not settled the controversy over the origin and phenotype of NK cells. On the other hand, NK cells do have specific properties not characteristic of other lymphocytes. In the mouse NK cells have been identified by a serological marker[7] and in man, although no unique surface marker has yet been defined, the more old fashioned, but not to be despised, use of cytological examination has revealed that most NK cells are large, granular lymphocytes with characteristic cytochemistry[8]. This is not to say that all cytotoxicity detected using cultured tumour lines *in vitro* as targets is due to this cell type. Certainly, activated lymphocytes which carry undoubted T-cell markers will lyse a variety of cultured target cells.

If many different cell types can lyse cultured tumour targets *in vitro* and if freshly isolated cells are in general insensitive to NK activity, what can be made of claims that NK cells operate *in vivo* as a surveillance mechanism against tumours? Several authors have shown that NK cells can mediate resistance *in vivo* against transplantable tumours. Many of these experiments are, however, difficult to interpret because NK cells cannot be purified. Kasai *et al.* (this issue of *Nature* p.334) have now adopted a rather different approach: they have used systemically administered antibody to asialo GM1 surface membrane glycolipid to block NK activity and shown a dramatic reduction in resistance to syngeneic tumour. The relevance of their results to natural tumour immunity is not entirely clear because they used a well established transplantable tumour cell line. The difficulty with such cell lines is that they generally contain tumour viruses and it may be against the viruses that the immune response is directed.

Another line of evidence that could be interpreted as indicating a role for NK cells in protection against virus infection is the observation that NK activity is increased by interferons[9]. Furthermore, antibody to interferon has been shown to block NK function in nude mice, and to allow the growth of xenogeneic cell lines even when these were persistently infected with viruses[10]. Such data are highly reminiscent of much older findings in mice immunosuppressed for long periods with anti-lymphocyte serum in which the tumours which arose were shown to be due to polyoma virus[11]. Taken together, these results may be more consistent with a view of the NK cells as a surveillance mechanism predominantly directed at viruses not tumours. This line of reasoning suggests that therapeutic manoeuvres, such as administration of interferon, which may activate NK cells, will be most effective against virus-induced tumours. In man, the successful use of interferon to treat laryngeal papilloma may be an example of this[12].

The increase in NK activity mediated by interferon results at least partly from recruitment of previously inactive cells. But interferon also affects target cells: they become more resistant to lysis and have been shown to express more major histocompatibility complex (MHC) antigen. This latter effect, while it may not account for the change in susceptibility to NK cells, might well make virus-infected cells more susceptible to T cell-mediated killing because of the importance of MHC antigens as target structures for T cells.

This raises the question of whether the function of NK cells need necessarily be the killing of tumour cells. Other data suggest a regulatory role: interferon may be produced by NK cells themselves, and play a part in the generation of suppressor T cells[13]. This implies that NK cells are an integral part of the immunological network of T and non-T cells. Since immunologically active cells are regulated by multiple signals (MHC antigens, specific antigens, helper and suppressor factors), it is not surprising that NK regulation has now been shown to depend on signals other than interferon: Henney *et al.* (this issue of *Nature*, p.335) have shown that their activity is augmented by interleukin-2. Whether this is a direct effect or mediated indirectly via another cell type is not established. It thus remains possible that the final mediators of the augmentation are interferons produced by intermediary cells induced by interleukin-2. This would imply synergistic effects of different interferons on NK cells. Irrespective of the details of the mechanism of the interleukin-2 effect however, this finding once again places the NK cell firmly in the immunological network.

In summary, recent data seem to be leading away from earlier and rather sterile arguments about the phenotype and effector specificity of NK cells (although these remain important unresolved questions) and towards a deeper understanding of the biology of NK cells *in vivo* and *in vitro*. The new tools available, monoclonal antisera[14], *in vitro* grown normal[15] and tumour[16] cell lines and purified mediators, should allow rapid progress.

Peter Beverley is ICRF Senior Scientist at the ICRF Human Tumour Immunology Group, University College Hospital Medical School, London.

1. Roder, J.C. & Haliotis, T. *Immun. Today* **1**, 96 (1980).
2. Herberman, R.B. *et al. J. Immun.* **121**, 304 (1978).
3. Roder, J.C. & Duwe, A *Nature* **278**, 451 (1979).
4. Hansson, M., Kiessling, R. & Andersson, B. *Eur. J. Immun.* **11**, 8 (1981).
5. Knight, R.A. & Fitzharris, P.F. *Br. J. Cancer* **42**, 243 (1981).
6. Roder, J.C. *et al. Nature* **284**, 553 (1980).
7. Glimcher, L., Shen, F.W. & Cantor, H. *J. exp. Med.* **145**, 1 (1977).
8. Saksela, E. *et al. Immun. Rev.* **44**, 71 (1979).
9. Trinchieri, G. & Santoli, D. *J. exp. Med.* **147**, 1314 (1978).
10. Reid, L.M. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 1171 (1981).
11. Simpson, E. & Nehlsen, S.L. *Clin. exp. Immun.* **9**, 79 (1971).
12. Dunnick, J. *J. infect. Dis.* **142**, 293 (1980).
13. Kadish, A.S. *et al. J. exp. Med.* **151**, 637 (1980).
14. Secker, D.S. & Burke, D.C. *Nature* **285**, 446 (1980).
15. Dennert, G. *Nature* **287**, 47 (1980).
16. Johnson, D.R. & Jondal, M. *Nature* **291**, 81 (1981).

0028-0836/81/220287-01$01.00

ARTICLES

Demise of the Lemaître cosmological models and objects at large redshift

David Tytler

University College London, Gower Street, London WC1E 6BT, UK

Recent uniform samples of QSO Lyα absorption lines make possible new cosmological tests out to large redshifts and the detection of superclustering out to redshift z ~ 3. One test involves the Lemaître models of the Universe, which predict large peaks in the number of objects per unit redshift, near the redshift of their characteristic quasistatic period. No such peaks are seen in the Lyα cloud distribution.

AMONG the various standard isotropic homogeneous Friedmann–Robertson–Walker cosmological models[1,2] with nonzero cosmological constant, the Lemaître models are the most distinctive because of their long quasistatic or coasting period during which the Universe scarcely expands at all (Fig. 1). Light passing through a coasting period which lasts a long time will traverse a large distance, and may intercept many absorbers for a corresponding small change of redshift z. Thus in Lemaître models one expects to observe a pronounced peak in the number of objects observed per unit redshift, near z_s, the redshift of the quasistatic period[3].

The Lemaître models of interest all have postive space curvature and a negative q_0. They are characterized by a positive cosmological constant Λ (repulsive force[2]), which is fractionally larger than the value Λ_c for an Einstein-static universe given by the relation:

$$\frac{\Lambda}{\Lambda_c} = 1 + \varepsilon = \frac{27\sigma_0^2(\sigma_0 - q_0)}{(3\sigma_0 - q_0 - 1)^3} \qquad (1)$$

where the dimensionless density parameter $\sigma_0 = \Omega_0/2$, and ε is a free parameter in the range $0 < \varepsilon \ll 1$. The expected number of absorbing objects per unit redshift along a line of sight is given by[4]:

$$\begin{aligned} N(z) &= cH_0^{-1}\phi(z)\Sigma(z)(1+z)^2|\mathrm{RAD}|^{-1/2}, \\ &\qquad \text{clouds observed per unit } z \\ \mathrm{RAD} &= (H(z)/H_0)^2 = 2\sigma_0(1+z)^3 + (1 + q_0 - 3\sigma_0) \\ &\quad \times (1+z)^2 + \sigma_0 - q_0 \end{aligned} \qquad (2)$$

where H_0 is the local ($z = 0$) value for the Hubble 'constant' $H(z)$, $\phi(z)$ is the space density of absorbing clouds per unit co-moving volume and the mean cloud cross-section at some critical column density is given by $\Sigma(z)$. If the clouds do not evolve then $N(z)$ depends only on the cosmological redshift–distance relation giving:

$$N_{\cos}(z) = N_0(1+z)^2|\mathrm{RAD}|^{-1/2} \qquad (3)$$

where N_0 is the local value of $N(z=0) = cH_0^{-1}\phi_0\Sigma_0$. The redshift z_p of a peak in $N(z)$ is determined by the values of σ_0, $\varepsilon(q_0)$ in equations (1) and (2). A large density σ_0 is required to give low z_p.

$N(z)$ peak test using galaxies and QSOs

Complete and unbiased samples of the redshift distribution of objects are required to search for $N(z)$ peaks. Several peaks have in fact been reported[5,6], but these are now considered to be purely the result of observational selection effects.

For $\sigma_0 > 1.14$ and small ε, peaks lie at $z_p < 0.5$ where galaxies are directly observable. Current surveys are only complete to bright magnitude limits (for example, $m_J \leqslant 15$ in ref. 7), and contain few galaxies with $z > 0.1$. In addition the redshift distribution is dominated by cosmologically small-scale ($\lesssim 20$ Mpc) galaxy clustering[7], making samples unsuitable for searches for Lemaître peaks.

To extend the test to Lemaître models with lower densities we must use surveys of objects at large redshift. For $\sigma_0 < 0.25$, $z_p \geqslant 1$ and for $\sigma_0 < 0.05$, $z_p \geqslant 2$, while at the lower limit for the density of the Universe based on observed galaxies Gunn and Tinsley[8] choose $\sigma_0 \geqslant 0.01$ giving $z_p \leqslant 3.8$. Only QSOs, other radio sources and probably X-ray sources are directly observable at these redshifts.

Large numbers of QSOs are now being discovered by low-dispersion spectroscopic surveys[9]. These surveys are strongly biased towards the detection of QSOs at emission redshifts for which at least two strong emission lines lie in the observational window and consequently show a strong peak in the number of QSOs at $z \sim 2$. This peak, and a narrow component in it at $z = 1.95$, have previously been regarded[6] as evidence for Lemaître models with $\sigma_0 \sim 0.05$. However, these peaks are adequately explained as selection effects[9,10], as is testified by their absence from samples of QSOs selected from radio surveys[11].

Aside from the above selection effects, all surveys of the distribution of objects in redshift which are based on the apparent observed magnitude of objects, in any waveband, will be strongly biased towards detecting fewer objects as redshift increases due to the apparent (not absolute) faintness of distant objects. Current samples of objects based on the flux emitted by each object are then inadequate for tests of the presence of Lemaître model $N(z)$ peaks[3]. It is in comparison with these data that QSO absorption lines provide a very powerful new data base for cosmological tests.

QSO absorption lines

Two major classes of intervening narrow line absorption clouds have now been distinguished in QSO spectra[4]: the metal line

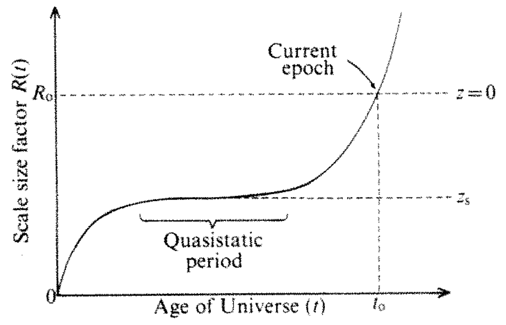


Fig. 1 Scale size of the Universe as a function of time since the Big Bang for a Lemaître model.

0028-0836/81/220289—04$01.00

systems, and the far more numerous Lyα clouds. It seems probable that a large fraction of the metal line systems arise in the outer halo regions of galaxies which happen to lie along the line of sight[12-14].

Working with an extremely heterogeneous sample of 136 metal line systems from 637 QSOs, Canuto and Owen[5] have found a broad peak in $N(z)$ near $z = 2$ which they interpret as evidence for a Lemaître model with $\sigma_0 = 0.05$, $\varepsilon \sim 10^{-3}$. In fact this peak probably also has the same origin as that in the distribution of the spectrally discovered QSOs themselves.

To minimize the extent of selection effects in samples of absorption line systems, Young *et al.*[13] propose strict criteria which each individual system must satisfy if it is to be included in a sample used to study redshift distributions. Seven QSOs for which results are available[4,15] yield a well defined sample of 15 systems from a total sample window of $1.5 < z_{\rm abs} < 3.3$. These systems are uniformly distributed in the redshift windows of their QSOs—showing no clumping at $z_{\rm abs} \sim 2$.

QSO Lyα absorption lines

Lyα line samples are seen as particularly valuable in view of the widespread occurrence of redshift dependent selection effects in other data. The properties of these absorption lines in QSO spectra have been extensively discussed by Sargent *et al.*[4], who conclude that the lines probably arise in intergalactic clouds. These lines are about an order of magnitude more numerous than the metal line systems.

Of fundamental importance is that, to first order, the detection of a Lyα cloud depends on only the intrinsic properties of that cloud, and not on its redshift (but see below regarding possible evolution of the clouds). Four main observational biases affecting the possible detection of a Lyα line have been discussed in detail elsewhere[4,15]. None of these closely interrelated difficulties is capable of either creating or hiding a Lemaître peak in the current data sample.

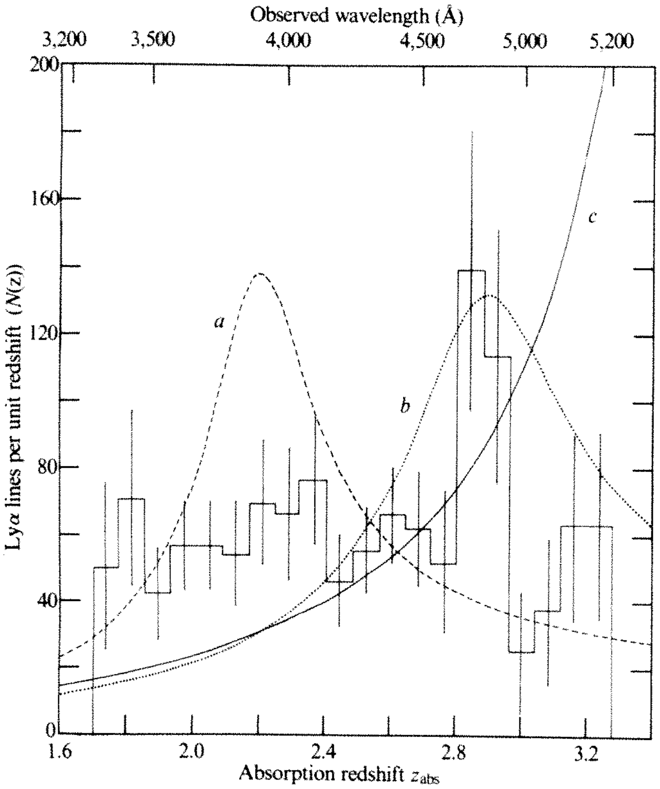


Fig. 2 The redshift distribution of Lyα lines. The statistical sample of 212 lines are shown in histogram form. The predicted $N_{\rm cos}(z)$ distributions for three Lemaître models are also shown. It is assumed that the Lyα clouds do not evolve with redshift. Model *a* (dashed curve) has $\log_{10}\sigma_0 = -1.4$ and $\log_{10}\varepsilon = -2.2$ giving a χ^2 probability of 9×10^{-6} for the observed distribution. Model *b* (dotted curve) for $\log\sigma_0 = -1.69$ and $\log\varepsilon = -2.0$ was chosen to provide a comparison with the data peak. The χ^2 probability is 10^{-9}. Model *c* (solid curve) has $\log\sigma_0 = -1.9$ and $\log\varepsilon = -2.2$ giving a peak at $z_{\rm p} = 3.5$, outside the observed range, and a probability of 10^{-14}.

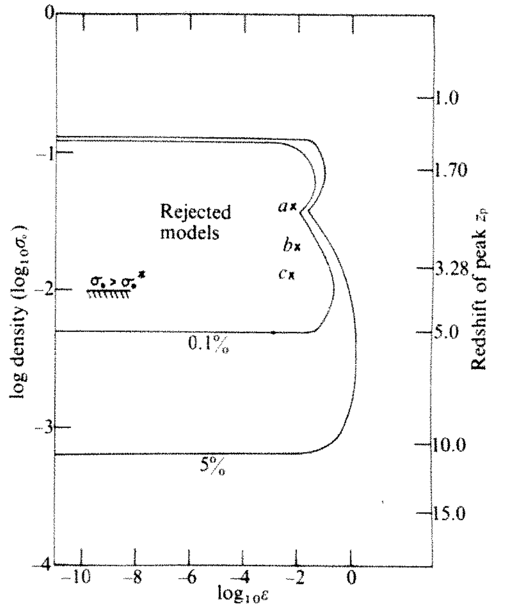


Fig. 3 Lemaître models which are rejected assuming that the Lyα clouds do not evolve. The $z_{\rm p}$ scale was obtained from the relation $\sigma_0^{-1} = (1+z_{\rm p})^3 - 3(1+z_{\rm p}) + 2$, which only applies in the limit $\varepsilon \to 0$. This scale gives $z_{\rm p}$ accurate to within 5% for models with $\log\varepsilon < -1$ and 0.5% for $\log\varepsilon < -2$. Models *a*, *b* and *c* from Fig. 2 are indicated. Percentage figures refer to the probability of falsely rejecting a model (that is, accepting that the observed distribution came, by chance, from the predicted one). Most of the models to the left centre of the 0.1% line are rejected with a negligible probability of false rejection. The small dip in the probability contours is purely the result of the geometry of fitting the predicted distributions to a data set of finite extent in redshift. The estimated lower limit on the density of the Universe[8] is shown by σ_0^*. The upper density limit is extremely uncertain, although $\log\sigma_0 < +1$ does seem safe.

The redshift distribution of the statistical sample of 212 Lyα lines from seven QSOs[4,15] is shown in Fig. 2. Data from more than one QSO (line of sight) contribute to all except for the first bin and the last six bins. The overall distribution is well represented by a constant density of 61±4 lines per unit z (1,216 Å in the observed frame). There is, however, a prominent peak at $z \sim 2.9$. The two peaked bins span the wavelength range 4,627–4,819 Å and contain 20 undoubtably real lines (P. J. Young, personal communication), all from the line of sight towards PKS2126−158. As the expected mean content for any two bins is ~9.6±0.6 lines, the feature is of marginal significance. A preliminary analysis of observations of five other QSOs (R. F. Carswell, personal communication and ref. 16) shows that line density enhancement at $z \sim 2.9$ occurs only towards PKS2126−158, and so it cannot be associated with a Lemaître model peak. A possible physical origin for the enhancement is discussed below.

Rejection of Lemaître models

The test is straightforward, involving a numerical evaluation of the probability that the observed Lyα line distribution arose from the distribution predicted by various cosmological models. The 212 lines were divided into 20 bins each of width $\Delta z = 0.079$, spanning a total redshift range of 1.70–3.28, and the predicted $N_{\rm cos}(z)$ distribution was integrated numerically across each bin. The normalization constant N_0 was estimated and the probablility that the observed distribution came from the predicted one can be found using the χ^2 distribution. The predicted $N_{\cos}(z)$ distributions for three particular models are shown in continuous form (before binning) in Fig. 2. Only models with $-1.83 < \log_{10}\sigma_0 < -1.13$ have peaks in the observed redshift range. Outlying models such as c ($\log_{10}\sigma_0 = -1.9$) are rejected because of a large variation of $N_{\cos}(z)$ in the observed z range. Note that Lemaître peaks are too broad to fit the observed data peak. Narrower peaks can be obtained by decreasing ε, but these are then too high. The complete set of Lemaître models which are rejected on the assumption of a non-evolving population of Lyα lines is shown in Fig. 3.

Evolution of the Lyα clouds

The observed distribution of Lyα clouds is the product of the cosmologically specific redshift–path-length relation $N_{\cos}(z)$, the cloud space density $\phi(z)$ and the mean cloud cross-section $\Sigma(z)$ at the critical H I column density. While the simplest assumption is that the observed flat distribution arises from the product of three slowly changing functions; we should consider the possiblity that $N_{\cos}(z)$, $\phi(z)$ and $\Sigma(z)$ all vary rapidly with redshift in a closely correlated manner such that their product gives the observed flat distribution.

Note that the peaks in $N_{\cos}(z)$ arise because the Universe spends a long time at an approximately constant scale size $R = R_0/(1+z_s)$. There is ample time for a significant amount of evoution in $\phi(z)\Sigma(z)$ as z changes slightly on passing through a Lemaître peak.

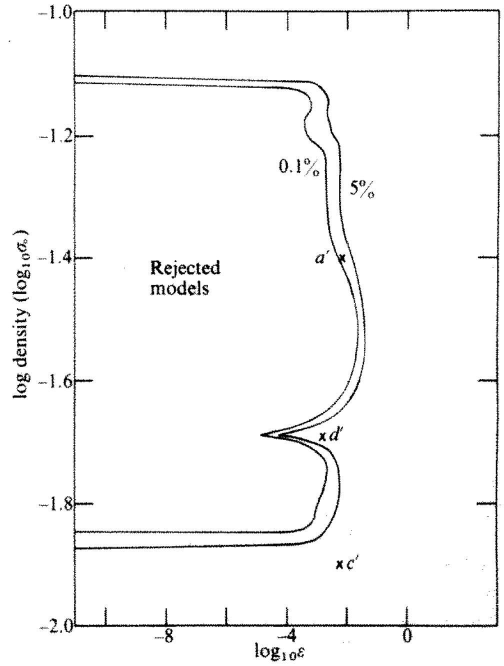


Fig. 5 The set of Lemaître models which are rejected on the assumption that the Lyα lines evolve as $(1+z)^G$. Models a', c' and d' from Fig. 4 are indicated. Percentage figures refer to the probability of false rejection of a model. The small dip in the proability contours near $\log\sigma_0 \sim -1.2$ arises from the geometry of fitting the theoretical distributions to a data set of finite range in redshift.

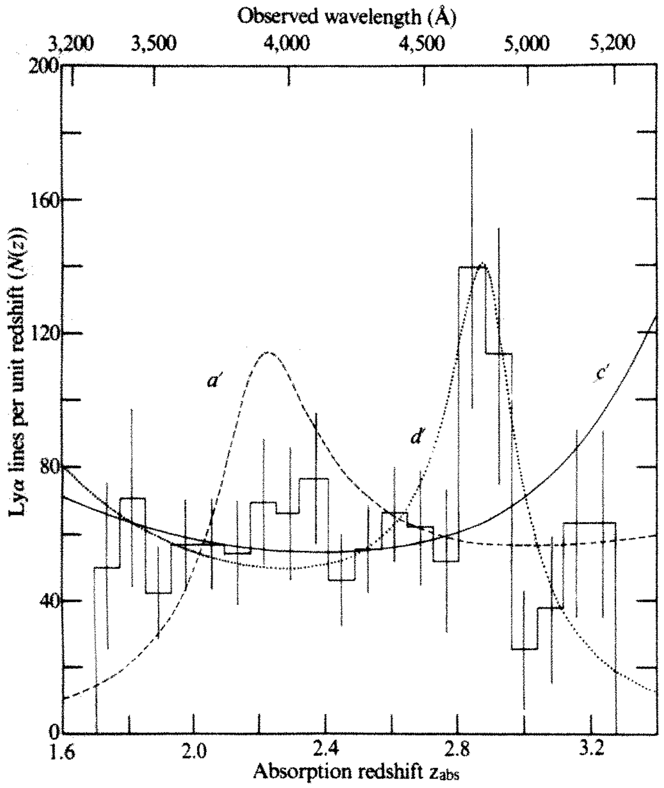


Fig. 4 The predicted distribution $N(z, G)$ of an evolving population of Lyα lines for three Lemaître models. The histogram shows the data, as in Fig. 2. Models a' (dashed curve) and c' (continuous curve) are the equivalents of models a and c in Fig. 2. Model a' has $G = 5.0$ and a χ^2 probability of 0.3%. Model c' has $G = -2.7$ and a probability of 28%. The third model d' has $\log\sigma_0 = -1.69$ and $\log\varepsilon = -2.8$ giving $G = -5.0$ and a χ^2 probability of 27%.

A particularly simple model for the evolution of Lyα clouds has been presented Sargent *et al.*[4] They propose that the clouds are pressure supported by an intergalactic medium which expands adiabatically with the Universe as a whole (see refs 17, 18 for alternative proposals). Adopting several simplifying assumptions, the resulting evolution is of the form $\phi(z)\Sigma(z) \propto (1+z)^G$, giving a predicted cloud distribution:

$$N(z, G) = N_0(1+z)^{2+G}|\mathrm{RAD}|^{-1/2} \qquad (4)$$

We cannot assign a unique value to the evolutionary parameter G without first accepting that some particular cosmological model is valid. Sargent *et al.* choose the standard Friedmann models ($\Lambda = 0$) and find that $G \sim -0.7$ is consistent with both the observed $N(z)$ as well as the line equivalent width distribution $n(W)$. In fact there is an error in the exponents of their equations (69)–(71); on correction we find that there is no longer any value of G which satisfies both $N(z)$ and $n(W)$. The obvious recovery is to drop the weakest of the many assumptions: that all the clouds have the same total mass. The result is a cloud mass distribution $n(M) \propto M^{-1}$ for $G \sim 0$ and $\sigma_0 = q_0 = \frac{1}{2}$; but G is still dependent on the *a priori* acceptance of a particular cosmological model (work in preparation).

In view of these uncertainties we treat G as a free parameter. We can then choose the values for G which make $N(z, G)$ in closest agreement with the observed Lyα distribution. The maximum likelihood method[4] was used to estimate G. Figure 4 shows the $N(z, G)$ distribution for two of the models for which $N_{\cos}(z)$ was shown in Fig. 2. A third model d' has the same σ_0 as b, but ε has been decreased to bring the $N(z, G)$ peak up to the height of the data peak. The complete set of models which are now rejected is shown in Fig. 5. The conspicuous dip in the probability contours at $\log_{10}\sigma_0 = -1.687$ arises because of the peak in the Lyα density at $z \sim 2.9$. As discussed above, it is highly unlikely that this feature is globally significant.

It is clear that the G parameter is able to counteract monotonic trends in $N_{\cos}(z)$ and produce more acceptable fits to the observed distribution. It cannot remove peaks. Most of the models with $\log_{10}\varepsilon < -2$ have $N_{\cos}(z)$ distributions which vary rapidly across the observed z range. Some of these models are accepted in Fig. 5, but only if we are prepared to allow large absolute values for $G(4 \leq |G| < 9)$. The metal line systems also seem to have an approximately flat $N(z)$ distribution. As it is unlikely that both the Lyα clouds and the distinctly different metal line systems both evolve in the same manner, it is improable that any model with $\log_{10}\sigma_0 < -1$ and $\log_{10}\varepsilon < -2$ is valid.

We conclude by outlining four general constraints which evolution of $\phi(z)\Sigma(z)$ would have to satisfy if a Lemaître peak were to be successfully disguised. (1) $\phi(z)\Sigma(z)$ must decrease as $N_{\cos}(z)$ increases on entering a quasistatic period, and then $\phi\Sigma$ must increase again as $N_{\cos}(z)$ decreases on leaving the quasistatic period. Two evolutionary processes would be required to reproduce the observed flat distribution of $N(z) \propto \phi(z)\Sigma(z)N_{\cos}(z)$. (2) Both evolutionary processes must be coupled to $N_{\cos}(z)$ and hence to the global dynamics of the Universe because $N_{\cos}(z) \propto H(z)^{-1}$. (3) $\phi\Sigma$ must return to close to its original value after the quasistatic period evolution—otherwise a sharp step would be seen in $N(z)$ at the redshift z_s. (4) Finally, if the evolution is to be able to disguise an arbitrarily sized peak, then the magnitude of the predicted change $\phi\Sigma$ must be related to the parameters which determine the peak height: ε or the duration of the quasistatic period. For small ε the required change in $\phi\Sigma$ during the quasistatic period is substantial: for $\log_{10}\varepsilon < -3.3$, $N_{\cos}(z_p)/N_{\cos}(z_p+0.5) \geqslant 10$.

It is clearly unlikely that more than one of these constraints can be met by any reasonable evolutionary model. Detailed models for the Lyα clouds, together with a better knowledge of the intergalactic medium and ionizing flux (see ref. 19) will be required to determine whether some of the weaker Lemaître peaks ($\log_{10}\varepsilon > -2$) could have been disguised by evolution; but it does seem safe to assume that strong peaks cannot have been entirely hidden. Figure 5 may be regarded as a minimal set of rejected models, with Fig. 3 specifying the maximal set for the current data.

Superclustering of Lyα lines

The peak in the density of Lyα lines in the line of sight towards PKS2126−158 raises the important possibility that Lyα lines show weak clustering on a very large scale[20]. Both the amplitude $(N-\bar{N})/\bar{N} \sim 1$ and the comoving size $\sim 60\,h_{100}^{-1}$ Mpc for the feature towards PKS2126−158 are similar to values found for local superclusters of galaxies[17,21].

First, we must explain why the overall apparent incidence and/or mean amplitude of superclustering of Lyα lines is low, as shown by the near zero correlation function found by Sargent *et al.* Geometrical factors will be important if Lyα superclusters display the flat or chain-like morphology of galaxy superclusters[17,22], and it can then be shown that only the occasional passage of a line of sight down the long axis of a supercluster could produce even a marginally detectable line density enhancement.

A fundamental question raised by Silk (personal communication) is: why should Lyα clouds follow the distribution of galaxies by showing superclustering, when they apparently do not show small scale galaxy clustering? For an explanation we must first account for the surprising lack of small scale clustering. Lyα clouds in a region of space containing galaxies can be expected to have a distribution which is clustered, following the dominant gravitational influence of the galaxies. The suggestion[4] that Lyα clouds are modified or destroyed in the vicinity of galaxies must be regarded with caution, because it is possible that a negative rather than zero correlation function would arise. We then arrive at the Sargent *et al.* hypothesis in which most Lyα clouds exist in intergalactic regions, in which there are few galaxies and consequently minimal clustering. Indeed large volumes of space containing few galaxies have been identified in three-dimensional maps of galaxy positions[22]. An important consequence of this model is that there may still be a fraction $<\frac{1}{3}$ of all observed Lyα clouds which do exist amongst galaxies and follow the galaxy clustering. The non-clustered clouds would provide sufficient dilution to produce a net correlation function which is almost zero.

These simple considerations show that great care will be required in the interpretation of Lyα cloud clustering. Jones[23] has stressed that it is difficult to think of observations which will help to distinguish between the rival 'graviational instability' and 'pancake' scenarios for the formation of galaxies. It can be expected that future studies of the evolution of clustering of absorption lines out to $z \sim 3$ will be of great relevance to this topic. For example, the existence of superclusters with density enhancements as large as ~ 1 at the early epoch of $z \sim 2.9$ would favour a 'pancake' scenario.

Conclusions

We have shown that, subject to the reasonable assumptions that the Lyα lines are at distances given by their redshifts[4,24–26] and that they do not evolve in a manner significantly more complex than a function of form $(1+z)^G$ over the redshift range 1.7–3.3, a wide range of Lemaître modes can be rejected beause of the lack of peaks in the observed distribution of Lyα lines.

This analysis is now being extended to models with higher densities in the range $0.1 < \sigma_0 < 10$ using a variety of data.

The distribution of an unbiased but small sample of narrow metal absorption line systems seen in QSO spectra shows no evidence for strong peaks in their redshift distribution. Clearly an earlier result showing a peak at $z_{\rm abs} \sim 2$ suffered from strong selection effects. The flat distribution supports the rejection of the Lemaître models as it is unlikely that the Lyα clouds and metal line clouds both evolve in the same manner, which would have to be closely correlated with the dynamics of the Universe as a whole, if the Lemaître model $N(z)$ peaks are to be extensively suppressed.

Evidence against the Lemaître models, including the $N(z)$ peak test used above, has been discussed elsewhere[5,6]. Tinsley[3] has pointed out that many of the other tests are subject to various difficulties and it now seems highly unlikely that any of them could be as strong or as direct as the $N(z)$ test using Lyα clouds for the low density (σ_0) Lemaître models shown in Fig. 3.

Further details of the present work here will be published elsewhere.

I thank A. Boksenberg, W. L. W. Sargent, P. J. Young, M. R. Rees and R. S. Ellis for many helpful discussions, and R. F. Carswell for permission to use unpublished data. J. Silk offered many helpful suggestions. Financial support from the SRC is acknowledged.

Received 14 January; accepted 24 April 1981.

1. Raychaudhuri, A. K. *Theoretical Cosmology* (Clarendon, Oxford, 1979).
2. Rindler, W. *Essential Relativity* 2nd edn (Springer, New York, 1977).
3. Tinsley, B. M. *Nature* **273,** 208–211 (1978).
4. Sargent, W. L. W., Young, P. J., Boksenberg, A. & Tytler, D. *Astrophys. J. Suppl.* **42,** 41–81 (1980).
5. Canuto, V. & Owen, *J. Astrophys. J.* **225,** 79–82 (1978).
6. Petrosian, V. *IAU Symp.* No. 63, 31–46 (1974).
7. Kirshner, R. P., Oemler, A. & Schechter, P. L. *Astr. J.* **83,** 1549–1563 (1978).
8. Gunn, J. E. & Tinsley, B. M. *Nature* **257,** 454–457 (1975).
9. Vaucher, B. G. & Weedman, D. W. *Astrophys. J.* **240,** 10–19 (1980).
10. Kjaergaard, P. *Phys. Scr.* **17,** 347 –351 (1978).
11. Wills, D. & Lynds, R. *Astrophys. J. Suppl.* **36,** 317–358 (1978).
12. Boksenberg, A. *Phys. Scr.* **17,** 205–214 (1978).
13. Young, P. J., Sargent, W. L. W., Boksenberg, A., Carswell, R. F. & Whelan, J. A. J. *Astrophys. J.* **229,** 891–908 (1979).
14. Sargent, W. L. W., Young, P. J., Boksenberg, A., Carswell, R. F. & Whelan, J. A. J. *Astrophys. J.* **230,** 49–67 (1979).
15. Carswell, R. F., Whelan, J. A. J., Smith, M. G., Boksenberg, A. & Tytler, D. *Mon. Not. R. astr. Soc.* (submitted).
16. Weymann, R. J. *Phil. Trans. R. Soc.* A **296,** 399–405 (1980).
17. Oort, J. H. *Astr. Astrophys.* (in the press).
18. Melott, A. L. *Astrophys.* **241,** 889–893 (1980).
19. Sherman, R. D. *Astrophys. J.* **240,** 737–743 (1980); *Astrophys. J.* (in the press).
20. Boksenberg, A. in *Radio Astronomy and Cosmology* (ed. Jauncy, D. L.) 193–215 (Reidel, Dordrecht, 1977).
21. Thuan, T. X. in *Physical Cosmology* (eds Balian, R., Audouze, J. & Schramm, D. N.) 277–289 (North-Holland, Amsterdam, 1980).
22. Einasto, J., Jôeveer, M. & Saar, E. *Mon. Not. R. astr. Soc.* **193,** 353–375 (1980).
23. Jones, B. J. T. *Phil. Trans. R. Soc.* A**296,** 289–298 (1980).
24. Wesson, P. S. *Gravity, Particles and Astrophysics* ((Reidel, Dordrecht, 1980).
25. Wills, D. *Nature,* **288,** 114–115 (1980).
26. Burbidge, G. *IAU Symp.* No. 92, 99–105 (1980); *Nature* **282,** 451–455 (1979).

Production of clones of homozygous diploid zebra fish (*Brachydanio rerio*)

George Streisinger, Charline Walker, Nancy Dower, Donna Knauber & Fred Singer

Institute of Molecular Biology, University of Oregon, Eugene, Oregon 97403, USA

Homozygous diploid zebra fish have been produced on a large scale by the application of simple physical treatments. Clones of homozygous fish have been produced from individual homozygotes. These clones and associated genetic methods will facilitate genetic analyses of this vertebrate.

HETEROZYGOSITY in diploid eukaryotes often makes genetic studies cumbersome. Here we describe the ready production of homozygotes in a vertebrate. Our methods have been developed to allow the eventual dissection of neuronal development by the use of mutant strains[1,2], but are also likely to find use in a wider range of applications.

We have chosen the tropical freshwater cyprinid zebra fish *Brachydanio rerio*[3,4] for several desirable attributes[5]: the generation time is only 3–4 months; at weekly intervals mature females lay several hundred eggs which develop rapidly and synchronously outside the mother; the fish are small (3 cm), hardy and easy to care for. Large-scale screening for mutants is possible because free-swimming 7-day-old fish exhibit many behavioural and morphological traits of the parent but are only a few millimetres long. Because development is normal between 25 and 31 °C (see also ref. 6) it should be possible to isolate temperature-sensitive mutants.

To produce homozygotes, eggs are activated by genetically impotent sperm. The haploid maternal set of chromosomes is allowed to replicate once, and partition of the duplicated chromosomes into two cells is prevented by hydrostatic pressure or heat shock. A cell with two identical sets of chromosomes is thereby produced. Our procedure is similar in principle to that described recently for the production of individual homozygous mice[7,8].

The generation of large numbers of homozygous diploids makes it possible to detect rare newly arisen recessive mutations. Linkage analysis is facilitated because the genotypes of gametes produced by a female are directly reflected by the phenotypes of her homozygous progeny. Large clones of diploid homozygous fish with identical genetic constitutions can be generated from individual homozygous females.

Production of homozygous diploids

The union of a sperm with an egg ordinarily serves both to activate it (initiate cleavage divisions) and to restore a diploid set of heterozygous chromosomes. We eliminated heterozygosity by activating zebra fish eggs with UV light-treated[9] sperm (UV sperm). These eggs develop into inviable haploid embryos with bent tails and vacuolated body cavities, as has been described for various lower vertebrates[10–12]. About 3×10^{-4} normal embryos appear; one such embryo whose karyotype has been examined was diploid and may have arisen through nondisjunction.

To generate viable embryos, haploid eggs were converted to diploids by either hydrostatic pressure or heat shock. In the late pressure treatment (LP), eggs activated by UV sperm are exposed to a combined pulse of ether and hydrostatic pressure at the first mitotic division (Fig. 1). Hydrostatic pressure dissociates microtubule subunits[13] and prevents cell division[14]; it has been used to inhibit the second meiotic division in frogs[15]. Ether inhibits cytokinesis in sea urchin eggs[16]. The establishment of the optimum treatment will be described elsewhere.

Females which produced grossly abnormal haploids (probably because of deleterious genes) were eliminated from our starting

Table 1 Survival of progeny from eggs treated to produce homozygotes

Population*	Treatment	No. of mothers	Fraction of fertile eggs which develop into normal-appearing 24-h embryos	Fraction of normal-appearing embryos which survive to: 9 Days with swim bladder†	Adulthood‡	Total no. of surviving progeny counted as adults or at 9 days with swim bladders
Starting	LP	90	0.29		0.20§	874
Selected I (8 different sets)¶	LP	49	0.20	0.73	0.68‖	831
Selected II						
(Set 1)	LP	4	0.28	0.69	0.61‖	146
(Set 2)	LP	17	0.09	0.46		73
Selected III (Set C29 × C32)	HS	28	0.33	0.61		501#
Clone C29 (Group III homozygote)	HS	18	0.25	0.38		85
Clone C29	EP	8	0.47	0.68		421
Clone C32 (Group III homozygote)	EP	8	0.51	0.81		494

* See Fig. 2.

† The fraction of embryos which develop swim bladders is influenced by the nutritional regime and general well-being of the mothers, especially for homozygous fish. Fish aged ⩾1 month are fed twice a day with a rotation of: Tetramin staple food (Tetra), Oregon test diet[28] and Oregon moist pellets (Bioproducts), liver-spinach pâté[29] and brine shrimp (*Artemia*) nauplii (eggs from Aquarium Products). If adult brine shrimps are available they can be substituted for the Oregon test diet, moist pellets and brine shrimp nauplii. Fish aged 4–14 days are fed paramecia and 9–30 day old fish are fed brine shrimp nauplii.

‡ Ordinarily, about 90% of the 9-day-old fish with swim bladders develop into adults.

§ Survival is for fish that have reached an age of 3 months.

¶ Each set is the product of a cross of two homozygous individuals [or in the case of selected population III (set C29 × C32), the product of crosses between two clones]. Different sets have different homozygous parents.

‖ Survival is for fish that have reached an age of 6 months.

Only a fraction of the progeny of each mother are reported here because some portions were subjected to other treatments.

0028-0836/81/220293—04$01.00

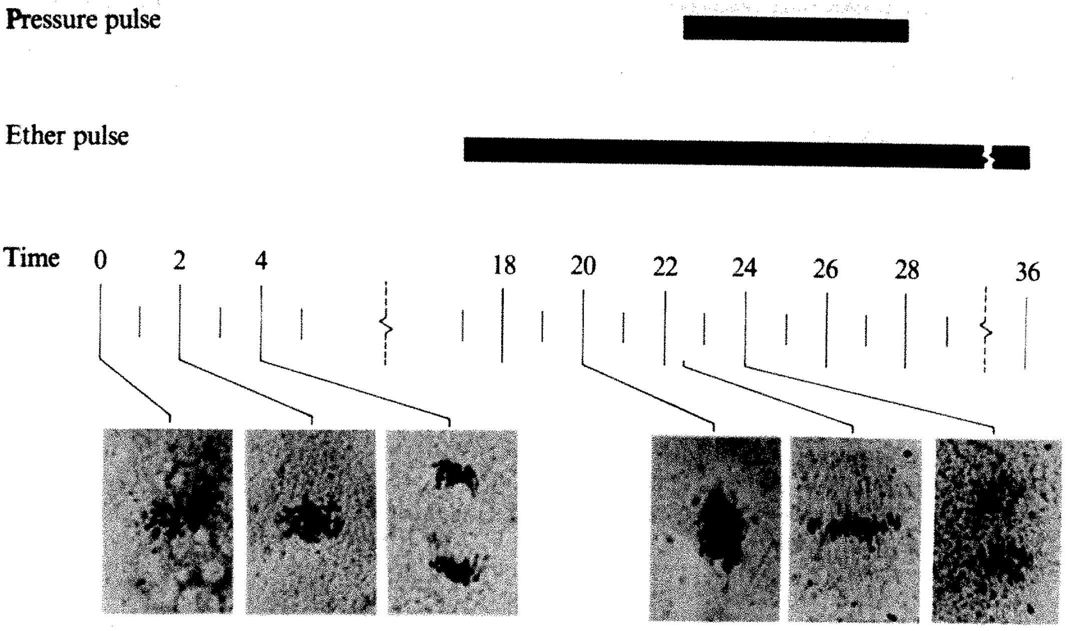


Fig. 1 Late pressure (LP) treatment. Groups of females were maintained at about 27 °C with a twofold excess of males, in a 14 h light–10 h dark cycle, at about one fish per l. To collect eggs, females were segregated from males about 17 h before the beginning of a light cycle. Shortly before fertilization was due, males were anaesthesized in 0.2 mg ml^{-1} ethyl *m*-aminobenzoate methanesulphonate (Finquel, Ayerst) [brought to *p*H 7.0 with Tris(hydroxymethyl)aminomethane], and sperm were expelled by gentle pressure with smooth forceps, and taken up into a capillary pipette. Sperm from four or five males were pooled in 0.4 ml of cold Hanks' saline to give a concentration of about 10^8 sperm per ml. Sperm in a watch glass were exposed to UV light from a 43-cm (Sylvania) germicidal tube at a distance of 29 cm for 2 min. For about 90 min after the beginning of the light cycle, eggs were expelled from anaesthesized females in 3.5-cm plastic Petri plates (Lux brand, as some batches of another brand proved toxic to zebra fish embryos) by gentle finger pressure. Eggs were activated by addition of 0.05 ml of sperm suspension and 0.5 ml water [glass-distilled and supplemented with 1 g l^{-1} sea salts (Instant Ocean, Aquarium Systems)] at 28.5 °C, followed by further water after about 30 s. Activated eggs were incubated at about 28.5 °C. Eggs were transferred to glass vials and at 17 min after activation water was removed and replaced by 2% v/v aqueous ether. The vials (filled to the brim) were closed by a thin sheet-rubber diaphragm, held in place with the vial's snap-on plastic cap into which a hole had been cut. The vial(s) were placed into a French press, and at 22.5 min the pressure in the chamber was raised to 8,000 p.s.i with a hydraulic press. (Pressure gauge dials on many hydraulic presses are calibrated to read pressures exerted on a 2-inch diameter piston, and readings must be converted to p.s.i.) At 28 min the pressure was reduced to ambient over 1 min and at 36 min the ether solution was gradually replaced by (salt-supplemented) water. The procedure was carried out at about 28.5 °C. Several hours after treatment, eggs were examined under a low-power microscope and any which appeared infertile or grossly abnormal were removed; the latter were incubated in separate dishes. The fraction of fertile eggs was determined in an aliquot which had not been exposed to LP; ordinarily, more than 80% of the eggs were fertile. The meiotic and mitotic figures shown under the time scale were from Feulgen-stained squashes of entire eggs. The first three inset figures show the second meiotic division and the next three the first mitotic division. The times indicated for the various stages are based on observations of a series of squashes. To obtain karyotypes 24–48 h (post-activation) embryos were removed from their shells and transferred to 10^{-2} M colchicine (Sigma) for 90 min. Their yolk sacs were then dissected away and the embryos transferred to 0.04 M sodium citrate and incubated for 7.5 min at 21 °C and for a further 7.5 min at about 4 °C. They were then transferred to a 3:1 (v/v) solution of methanol/acetic acid and stored at room temperature for about 30 min and then at about −20 °C. To make spreads, embryos were minced in a few drops of 50% acetic acid for 2 min, and then droplets of the suspension were repeatedly and rapidly deposited onto and removed from slides at 50 °C (ref. 30). After drying, the slides were incubated in Giemsa's stain, washed in water, dried and examined. For a mature fish, karyotypes from regenerating fins were prepared in a similar manner, except that the regenerating portions of fins were clipped[31] and incubated in medium containing 1.3×10^{-5} M colchicine, 10% fetal calf serum and 2×10^{-3} M glutamine. About eight spreads were counted for each fish.

population. About 29% of the LP-treated eggs from the starting population developed into embryos which appeared indistinguishable from normal diploid embryos at 24 h after fertilization (Table 1). At 24 h we examined karyotypes of 22 developed, but otherwise unselected embryos; 20 were diploid ($2N = 50$ chromosomes[17]), 1 was haploid and 1 aneuploid. The latter two embryos appeared abnormal and were not expected to survive. As the genetic contribution of sperm was eliminated by UV light, we expect that the diploid fish created by LP have two identical sets of maternally derived chromosomes and are thus homozygous.

About 20% of the normal-appearing 24-h LP embryos from the starting population survived to maturity (Table 1). The losses could be due to deleterious genes or to the destructiveness of the LP treatment. To select for populations of hardy fish we carried out the procedure outlined in Fig. 2. We crossed surviving group I homozygous fish which appeared relatively vigorous, long-lived and fecund. For crosses we used either spontaneously occurring homozygous males, or ones produced by the hormone treatments described below. The heterozygous offspring were raised to maturity. The LP treatment applied to eggs of these selected population I heterozygous fish yielded a new generation of homozygous fish, designated as group II in Fig. 2. As 68% of the normal-appearing group II homozygotes survived to maturity (Table 1), the enhanced survival is probably due to a decrease in the number of deleterious genes. Most mature LP fish are diploids: of 31 group II homozygotes, 30 had a modal number of 50 chromosomes in metaphase spreads obtained from colchicine-treated regenerating tail fins and one fish exhibited approximately equal numbers of spreads with 50 and 52 chromosomes. The most vigorous individuals among the group II homozygotes were crossed, and from the selected population II set produced by one pair of fish, the group III homozygotes were obtained by LP. From these fish we derived clones (for example C29 and C32) as described below.

In the heat shock (HS) treatment for generating diploids, eggs activated by UV sperm were incubated at an elevated temperature (for example, 41.4 °C) for a 2-min interval about 10 min before metaphase (at 13 min post-activation at ~28.5 °C). Heat shocks at analogous times inhibit a subsequent cycle of mitosis in *Physarum*[18], and have previously been used to inhibit the second meiotic division in lower vertebrates[19–21] and to produce tetraploid newts[22].

For many homozygous fish (and clones, see below) HS is far more effective than LP and for many populations 10–20% of HS-treated fertile eggs develop into adults (for example clone 29 and selected population III; Table 1). The karyotypes of 12 normal-appearing HS embryos were examined; all were diploids. The optimal temperature and the efficiency of the HS treatment depend on the genotype of the mother, as will be described elsewhere. For many strains the relative simplicity of the HS treatment makes it the method of choice.

The experiments described in the next section were designed to measure the extent of heterozygosity due to the occasional contribution of genes by UV sperm or to chromosome nondisjunction. In addition, we introduce some convenient genetic markers and illustrate analyses of gene–centromere distance and segregation.

Genetic contribution by UV sperm

The genetic contribution of the UV sperm was measured using a spontaneous recessive mutation (designated b1) in the *gol-1* (golden) gene. Fish that are *gol*+/*gol*+ or *gol-1*/*gol*+ become pigmented within 48 h of fertilization (Fig. 3*a*) but the *gol-1*/*gol-1* mutants lack pigmented melanocytes up to 76 h (Fig. 3*b*). The *gol-1* gene segregates as a typical mendelian recessive; 50% of the LP progeny of *gol-1*/*gol*+ heterozygotes are pigmented and 50% are pigmentless; crosses between heterozygous fish yield 25% progeny with the golden phenotype. Sperm of *gol*+/*gol*+ males were exposed to 10^3 erg mm^{-2} UV light and used to activate eggs of *gol-1*/*gol-1* females. Of more than 600 embryos, none was observed to have any pigment. As the UV sperm used for producing homozygous fish are exposed to a dose that is four times as high (4×10^3 erg mm^{-2}), heterozygosity due to the genetic contribution of the sperm seems unlikely.

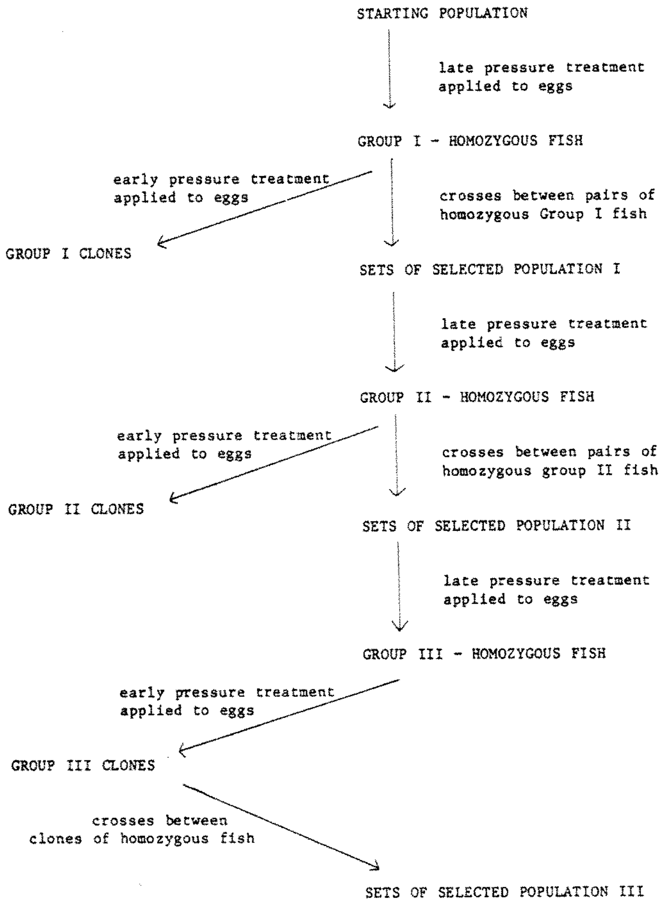

STARTING POPULATION
late pressure treatment applied to eggs
GROUP I - HOMOZYGOUS FISH
early pressure treatment applied to eggs
crosses between pairs of homozygous Group I fish
GROUP I CLONES
SETS OF SELECTED POPULATION I
late pressure treatment applied to eggs
GROUP II - HOMOZYGOUS FISH
early pressure treatment applied to eggs
crosses between pairs of homozygous group II fish
GROUP II CLONES
SETS OF SELECTED POPULATION II
late pressure treatment applied to eggs
GROUP III - HOMOZYGOUS FISH
early pressure treatment applied to eggs
GROUP III CLONES
crosses between clones of homozygous fish
SETS OF SELECTED POPULATION III

Fig. 2 Schematic diagram of the production of homozygous fish and clones.

Nondisjunction during meiosis I

Nondisjunction of the chromosomes of a heterozygous mother during meiosis will result in heterozygosity among her LP progeny. To estimate the magnitude of this contribution, we used electrophoretically different allozymes[23] which produce distinguishable phenotypes in heterozygotes and homozygotes.

Esterases have been described in zebra fish[24]; we found a set of three electrophoretically different esterase bands (designated as 3, 4 and 5) which seem to be specified by three alleles of a single gene we call *est-3*. On electrophoresis of individual fish extracts, one or two but never all three bands were observed. Phenotypes of homozygous and heterozygous fish will be described as 3, 4 or 5, and 3/4, 3/5 or 4/5, respectively. Another band of esterase activity (band 2) seems to be specified by a gene closely linked to *est-3*, as will be described elsewhere.

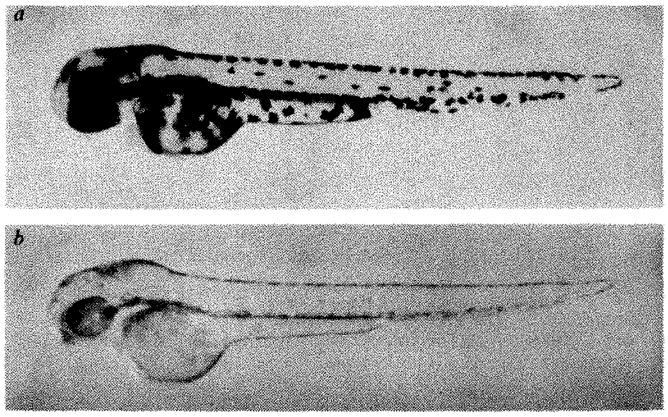


Fig. 3 Pigment differences between normal and mutant (*gol-1*) fish. *a*, A 3-day-old *gol*+/*gol*+ zebra fish; *b*, a 3-day-old *gol-1*/*gol-1* zebra fish. These fish are about 3.8 mm long.

To detect nondisjunction, we measured the frequency of heterozygotes among the LP progeny of heterozygous mothers who were formed by crossing pairs of homozygotes and exhibited 3/4, 3/5 or 4/5 band patterns. As shown in Table 2 and Fig. 4, no heterozygotes were observed among the 476 LP progeny examined. These results show that the three bands are specified by alleles or by very closely linked genes (we demonstrate below that recombination does occur in zebra fish).

Nondisjunction at the first meiotic division would result in heterozygosity in almost every case in which the unseparated chromosomes are retained in the egg. As no heterozygotes were observed among 476 progeny, heterozygosity for markers on the *est-3* chromosome occurs with a frequency of less than 1% (considering 95% confidence limits) among mature fish. Nondisjunction could, of course, be more frequent but result in lethality or it could be more frequent for other chromosomes.

Nondisjunction during meiosis II

Nondisjunction during the second meiotic division does not necessarily result in heterozygosity: the sister chromatids remaining in the egg before the second meiotic division will be identical except for regions separated from their original centromere by genetic recombination. Nondisjunction can be detected only in eggs that contain sister chromatids which are heterozygous.

To measure the frequency of heterozygous sister chromatids, eggs of 3/4, 3/5 and 4/5 mothers were subjected to a treatment which inhibits the second meiotic division[9,25,26]; in zebra fish, as in many other vertebrates, this division occurs after fertilization. Exposure of zebra fish eggs activated by UV sperm to a pulse of 8,000 p.s.i. hydrostatic pressure for 1.5–6 min post-activation (the early pressure or EP treatment) effectively inhibits the second meiotic division and results in >30% viable diploid embryos (Table 1). These fish contain a diploid complement of chromosomes because the sister chromatids are retained in the

Table 2 Segregation of band pattern phenotypes among LP progeny of heterozygous parents

Mothers with bands	No. of LP progeny with bands			
3 and 4	(3)	(4)	(3 and 4)	Total
	72	59	0	131
3 and 5	(3)	(5)	(3 and 5)	Total
	73	58	0	131
4 and 5	(4)	(5)	(4 and 5)	Total
	105	109	0	214
		Total no. of LP progeny:		476

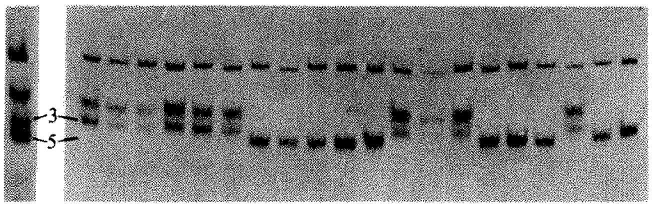


Fig. 4 Electrophoretic patterns. Fins of appropriate fish were clipped, stored frozen in a small volume of buffer overnight, and after thawing, aliquots of the supernatant were applied to 10% acrylamide slab gels. After electrophoresis, gels were incubated with the esterase substrate α-napthyl acetate and then stained with the dye Fast Blue RR[32]. Electrophoretic patterns are shown for a heterozygous (3/5) female (in the first lane) and for a group of LP progeny of this female (in the remaining 20 lanes). Note the presence of either band 3 or band 5 in the lanes of progeny extracts. A gene specifying the band 2 esterase co-segregates with the *est-3* gene.

egg nucleus. Among the progeny fish which developed from these sets of EP-treated eggs, 14% (87/626) exhibited two of the appropriate bands and were thus heterozygous. These results and other gene–centromere distances will be described more fully elsewhere.

Nondisjunction at the second meiotic division could thus have been detected in only 14% (or about 67) of the 476 LP progeny described in Table 2. The absence of heterozygotes suggests that fewer than 6% of mature fish (based on 0/67, with 95% confidence limits) develop from eggs which have experienced second meiotic division nondisjunction of the chromosome bearing the *est-3* gene.

Clones of homozygous diploids

For a ready interpretation of the consequences of particular mutations, standard and mutant strains ideally should differ at only one site. We generated clones of homozygous fish by EP or HS treatment of eggs of homozygous mothers. The many fish in a particular clone are all derived from the same single cell (the UV sperm-activated egg from which the clone's mother developed) without sexually contributive processes.

Clones have also been propagated by matings of 'hormone-created' males[27] and their normal sisters. Ordinarily, more than 95% of the fish in the C29 clone are females. Exposure of C29 embryos for 2 weeks, beginning 1 day after fertilization, to 1.7×10^{-7} M 17α-methyltestosterone (Sigma) resulted in the production of more than 75% males (defined by production of functional sperm) and few, if any, females.

The clones derived from various homozygous individuals of selected population II (Fig. 2) differed with respect to viability, longevity and fecundity. Some clones contained a small fraction of individuals with characteristic physiological abnormalities, probably because of the presence of recessive, deleterious, low-penetrance alleles inherited from the starting population. Such clones as C29 and C32 (group III) are similar to the starting population with respect to desirable properties. Indeed, successive generations of these clones have been in routine use in our laboratory for large-scale experiments. However, although suitable for routine use, C29 and C32 fish seem more subject to stress, exhibit more stringent dietary requirements and produce a larger fraction of poorly developing eggs than do wild-type fish.

More vigorous hybrid fish were produced by crossing two clones of homozygous fish (for example, C29 and C32 to produce selected population III). These hybrids are satisfactory for the isolation of mutants because they are free of lethal genes. As several of the matings described in Fig. 2 are between siblings, the progeny of selected population III are expected to be homozygous at about 75% of the genome.

A surprising observation is that some clones (derived by HS or EP from single homozygous females) are predominantly male. Subsequent generations of such clones also contain a high frequency of males. The genetic basis of sex determination in zebra fish is unknown.

Conclusions

The large-scale production of homozygous zebra fish has already enabled us to isolate mutants, to measure frequencies of induced recessive lethals, to recognize fish heterozygous for recessive lethals and to measure recombination frequencies.

At present, heat shock seems to be the simplest method for producing homozygotes, although for some clones we have not been able to find productive temperature regimes.

The genetic methods we describe may be applicable to fish such as trout, salmon and catfish, which serve as human food. Treatments which inhibit the second meiotic division have been used successfully for such fish[22,23], and G. Thorgaard (personal communication) has been able to inhibit the first cleavage division in a fraction of treated trout eggs with heat shocks. The generation of clones may facilitate isolation of stocks with desired characteristics. In analogy with common agricultural practice, it may be possible to maintain clones for the generation of high-yield hybrid lines for harvest. If such endeavours are contemplated, it is crucial to keep in mind the need for preserving large fractions of the natural populations of any given species in their natural environment. As clones such as we have developed contain only a small fraction of the gene pool of the natural population, a large-scale displacement of natural populations by specific clones could easily lead to losses of entire species due to unexpected selective forces.

We thank Harriet V. Kuhnlein for her assistance at the beginning of this project, S. Chakrabarti and J. Golin for providing unpublished data, our colleagues at the Institute of Molecular Biology and Department of Biology for encouragement, Ira Herskowitz for enthusiastic interest and suggestions concerning this manuscript, R. O. Sinnhuber and J. D Hendricks for Oregon moist pellets and test diet, Dr G. Thorgaard for advice in the preparation of karyotypes, and Kathleen Teichman and Julie Dunn for secretarial help. This work was supported in part by NSF grant P3B 3128 and NIH grant R01 GM 22731.

Received 25 August 1980; accepted 13 March 1981.

1. Benzer, S. *Proc. natn. Acad. Sci. U.S.A.* **58**, 1112–1119 (1967).
2. Brenner, S. *Genetics* **77**, 71–94 (1974).
3. Hamilton, F. *Account of the Fishes Found in the River Ganges and its Branches* (Edinburgh, 1822).
4. Buchanan, F. *Rev. Bull. Gen. Universel (Fereissac)* **3**, 253 (1823).
5. Laale, H. W. *J. Fish. Biol.* **10**, 121–173 (1977).
6. Schirone, R. C. & Gross, L. *J. exp. Zool.* **169**, 43–52 (1968).
7. Hoppe, P. C. & Illmensee, K. *Proc. natn. Acad. Sci U.S.A.* **74**, 5657–5661 (1977).
8. Markert, C. L. & Petters, R. M. *J. exp. Zool.* **201**, 295–302 (1977).
9. Nace, G. W., Richards, C. M. & Asher, J. H. Jr *Genetics* **66**, 349–368 (1970).
10. Hertwig, G. *Arch. mikrosk. Anat. EntwMech.* **77**, 165–209 (1911).
11. Fankhauser, G. *J. Hered.* **28**, 3–15 (1937).
12. Subtelny, S. *J. exp. Zool.* **139**, 263–305 (1958).
13. Tilney, L. G., Hiramoto, Y. & Marsland, D. *J. Cell Biol.* **29**, 77–95 (1966).
14. Marsland, D. *Int. Rev. Cytol.* **5**, 199–227 (1956).
15. Dasgupta, S. *J. exp. Zool.* **151**, 105–121 (1962).
16. Swann, M. M. *Expl Cell Res.* **7**, 505–517 (1954).
17. Endo, A. & Ingalls, T. H. *J. Hered.* **59**, 382–384 (1968).
18. Brewer, E. N. & Rusch, H. P. *Expl Cell. Res.* **49**, 79–86 (1968).
19. Fankhauser, G. & Watson, R. C. *Proc. natn. Acad. Sci. U.S.A.* **28**, 436–440 (1942).
20. Swarup, H. *J. Genet.* **56**, 129–142 (1959).
21. Purdom, C. E. *Heredity* **29**, 11–24 (1972).
22. Fischberg, M. *J. Embryol. exp. Morph.* **6**, 393–402 (1958).
23. Holmes, R. S. & Whitt, G. S. *Biochem. Genet.* **4**, 471–480 (1970).
24. Hart, N. H. & Cook, M. *Comp. Biochem. Physiol.* **B54**, 357–364 (1976).
25. Lindsley, D. L., Fankhauser, G. & Humphrey, R. R. *Genetics* **41**, 58–66 (1956).
26. Volpe, E. P. & Dasgupta, S. *J. exp. Zool.* **151**, 287–301 (1962).
27. Yamamoto, T. in *Fish Physiology* Vol. 3 (eds Hoar, W. S. & Randall, D. J.) 117–175 (Academic, New York, 1969).
28. Sinnhuber, R. O., Hendricks, J. D., Wales, J. H. & Putnam, G. B. *Ann. N. Y. Acad. Sci.* **298**, 389–408 (1977).
29. Gordon, M. in *The Care and Breeding of Laboratory Animals* (ed. Farris, E.) 345–449 (Wiley, New York, 1950).
30. Kligerman, A. D. & Bloom, S. E. *J. Fish. Res. Bd. Can.* **34**, 266–269 (1977).
31. Scheel, J. J. *Z. zool. Syst. Evolut.-forsch.* **10**, 180–209 (1972).
32. Shaw, C. R. & Prasad, R. *Biochem. Genet.* **4**, 297–320 (1970).

DNA-dependent RNA polymerase II of plant origin transcribes viroid RNA into full-length copies

Hans-Richard Rackwitz, Wolfgang Rohde & Heinz L. Sänger

Max-Planck-Institut für Biochemie, D-8033 Martinsried bei München, FRG
Arbeitsgruppe Pflanzenvirologie, Justus-Liebig-Universität, D-6300 Giessen, FRG

DNA-dependent RNA polymerase II purified from healthy plant tissue is capable of synthesizing linear (−)-viroid RNA copies of full length from (+)-viroid RNA templates in vitro. *Together with the specific α-amanitin sensitivity of viroid replication observed* in vivo, *these findings suggest that viroids replicate by an entirely novel mechanism in which infecting viroid RNA molecules are copied by the host enzyme which is normally responsible for the synthesis of nuclear precursors to messenger RNA.*

VIROIDS, the smallest known agents of infectious diseases, are responsible for several economically important diseases of higher plants. They are single-stranded covalently closed circular RNA molecules consisting of about 360 nucleotides, which form a highly base-paired rod-like structure with unique properties[1-3]. Thus, the 359 nucleotides of the potato spindle tuber viroid (PSTV)[2] are arranged in 26 double-stranded segments separated by 25 regions of unpaired bases embodied in single-stranded internal loops; there is a loop at each end of the rod-like molecule. These features provide the viroid RNA molecule with structural, thermodynamic and kinetic properties very similar to those of a double-stranded DNA molecule of the same molecular weight and G · C content[3].

The mechanism of viroid replication has long been puzzling. The limited coding capacity of the viroid genome argued against the synthesis of a viroid-specific replicase in the infected cell and, indeed, viroid-specific polypeptides have not been found. Thus, viroid replication would depend on the host nucleic acid-synthesizing machinery. Nor do viroids exist in a DNA form integrated into the host genome and from which new progeny viroid could be transcribed[4-6], since stringent molecular hybridization has shown that this is not the case[7,8,30].

Our finding that viroid synthesis is inhibited *in vivo* by an intracellular concentration of 10^{-8} M α-amanitin strongly suggested that the nuclear DNA-dependent RNA polymerase II is involved in the initial stages of viroid replication[9]. We now report that this enzyme, purified from healthy viroid host plant tomato tissue and from wheat germ, is able to synthesize full-length viroid RNA copies *in vitro* with viroid (+)RNA as template. (The isolated and infectious viroid RNA is arbitrarily called (+) viroid RNA.)

Template specificity of RNA polymerase II

DNA-dependent RNA polymerase II was isolated from either callus or green leaf tissue of tomato (*Lycopersicon peruvianum* (L.)Mill.), an excellent host for different viroid 'species'. The purification followed essentially the procedure developed by Jendrisak and Burgess[10] for wheat-germ polymerase II (Table 1). With both enzymes, standard single-stranded DNA templates are transcribed less efficiently than double-stranded

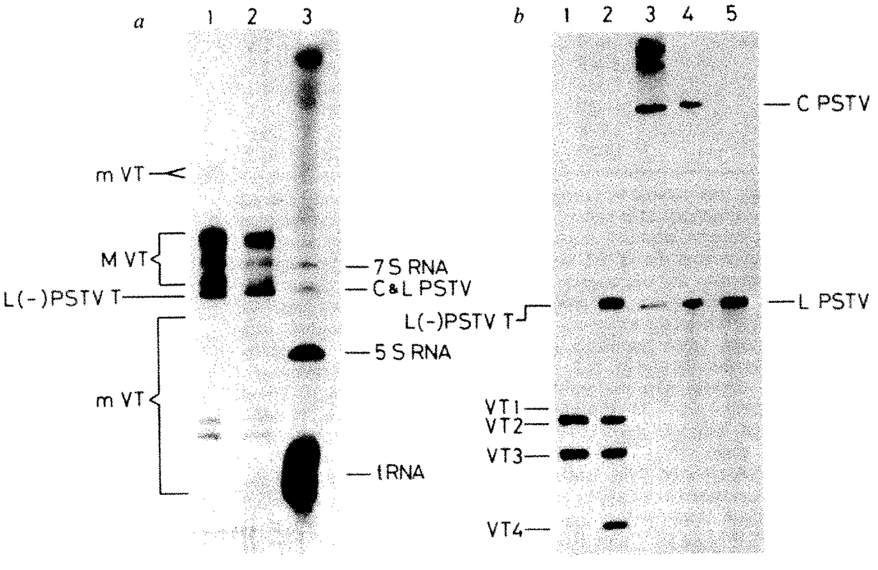


Fig. 1 Analysis by polyacrylamide gel electrophoresis (PAGE) and autoradiography of ^{32}P-labelled RNAs transcribed from viroid (PSTV) RNA by DNA-dependent RNA polymerase II from tomato tissue and wheat germ. Incubation mixtures for RNA synthesis contained at a final volume of 500 μl the cofactors described in Table 1 legend as well as 500 ng of purified PSTV RNA, 1.5 μg of tomato and wheat-germ RNA polymerase II, respectively, and 100 μCi of each of the [α-^{32}P]-labelled ribonucleoside triphosphate (specific activity 400 Ci mmol^{-1}). The reaction was stopped after 4 h at 37 °C by adding EDTA to 0.01 M. After the addition of 100 μg of yeast carrier tRNA and phenol extraction, the mixtures were chromatographed on 20-ml columns of Sephadex G-50 in (TNE) buffer (10 mM Tris-HCl *p*H 7.4, 100 mM NaCl, 1 mM EDTA). The peak fractions of radioactively labelled RNA in the exclusion volume of the columns were pooled, RNA precipitated with 3 vol of ethanol and collected for PAGE by centrifugation. Pellets were redissolved in 20 μl of 89 mM and 22.3 mM TEB buffer, respectively, containing 8 M urea and xylene cyanol FF and bromophenol blue as dye markers. Marker RNAs were ^{32}P-labelled, LiCl-soluble RNA from PSTV-producing potato callus isolated according to ref. 27 and ^{125}I-labelled PSTV prepared according to the method of Commerford[29]. *a*, PAGE on a 5% polyacrylamide slab gel (140 × 190 × 1.5 mm) at 10 °C and in the presence of high salt (89 mM) TEB buffer and 8 M urea. 1, Transcriptional products of wheat-germ RNA polymerase II. 2, Transcriptional products of tomato RNA polymerase II. 3, ^{32}P-labelled cellular RNA from PSTV-infected tomato callus[27]. MVT, major viroid (PSTV) transcripts; mVT, minor viroid (PSTV) transcripts; L (−) PSTV T, linear viroid (PSTV) transcript. *b*, PAGE on a 5% polyacrylamide gel at 40 °C in the presence of low salt (22.3 mM) TEB buffer and 8 M urea. 1, Transcriptional products of wheat-germ enzyme. 2, Transcriptional products of tomato enzyme. 3, ^{32}P-labelled cellular RNA. 4, Circular (C PSTV) and linear (L PSTV) ^{125}I-labelled PSTV RNA. 5, Linear ^{125}I-labelled PSTV RNA. VT1-VT4, viroid-(PSTV) RNA-directed RNA transcripts.

0028-0836/81/220297—05$01.00

DNA templates but the transcription of single-stranded DNA templates is considerably improved in the presence of oligo-RNA primers (Table 1).

Synthetic and natural RNAs also served as templates for transcription, thus confirming previous results on the same enzyme from pea and cauliflower[11-13]. In our case, tobacco mosaic virus (TMV) RNA and soluble RNA (5S RNA and tRNA) from yeast were relatively poor templates, whereas the synthetic template/primer poly(rA)/oligo(rU) is transcribed with high efficiency. Most surprisingly, however, of all natural RNA templates tested, viroid RNA is transcribed with the highest efficiency by DNA-dependent RNA polymerase II from tomato and wheat germ (Table 1). When PSTV, citrus exocortis viroid (CEV) and chrysanthemum stunt viroid (CSV) were compared with respect to their template efficiency, PSTV RNA was transcribed best by the tomato polymerase II. PSTV was propagated in and isolated from tomato (*Lycopersicon esculentum* Mill. cv. 'Rentita') whereas for CEV *Gynura aurantiaca* DC and for CSV *Chrysanthemum morifolium* Ramat. cv. 'Yellow Delaware' were used as host plants. These three viroids also showed marked differences of their oligonucleotide fingerprints and hence their primary sequence[14]. Although overall transcription was about the same when circular and linear viroid PSTV RNA molecules were used as templates, the pattern of product formation varied slightly between different template preparations.

Table 1 Transcription activity of different DNA and RNA templates for DNA-dependent RNA polymerase II from tomato tissue and wheat germ

		Incorporation			
		Tomato polymerase		Wheat-germ polymerase	
a DNA templates	^{3}H-NMP	Counts per 30 min	%	Counts per 30 min	%
Poly(dA)	UMP	2,425	13.2	1,583	8.3
Poly(dA) · poly(dT)	UMP	6.175	33.7	11,514	62.3
Poly(dA-dT)	UMP	18,312	100	18,470	100
Poly(dG-dC)	CMP	15,240	83.2	6,488	35.1
Poly(dA) · oligo(rU)	UMP	30,160	164.7	19,225	104.2
Tomato DNA	UMP	4,210	23.0	1,849	10.0
Salmon sperm DNA	UMP	5,305	28.9	2,030	11.0
Calf thymus DNA	UMP	5,460	29.8	2,476	13.4
b RNA templates		Counts per 60 min	%	Counts per 60 min	%
Poly(rA)	UMP	560	4.4	257	1.4
Poly(rA) · poly(rU)	UMP	256	2.0	316	1.7
Poly(rA-U)	UMP	366	2.9	739	4.0
Poly(rA) · oligo(rU)	UMP	12,630	100	18,287	100
TMV RNA	UMP	609	4.8	640	3.5
Yeast soluble RNA	UMP	220	1.7	358	1.9
PSTV RNA	UMP	1,658	13.1	1,366	7.5
CEV RNA	UMP	1,038	8.2	1,108	6.1
CSV RNA	UMP	1,196	9.5	1,220	6.7
L PSTV RNA	UMP	1,315	10.4	1,201	6.6
C PSTV RNA	UMP	1,463	11.6	1,135	6.2

RNA polymerase was isolated by precipitation from the clarified crude extract by Polymin P, re-extraction of the pellet by high salt buffer, ammonium sulphate precipitation and column chromatography on DEAE-cellulose. Tomato polymerase II was eluted from the DEAE column with 0.4 M ammonium sulphate buffer, *p*H 7.9, whereas the wheat-germ enzyme elutes at a 0.25 M salt step. According to its subunit pattern in SDS-containing polyacrylamide gels, the tomato enzyme differs from the wheat-germ enzyme, but it behaves similarly with respect to substrates, DNA templates, metal ions and specific inhibition by 10^{-8} M α-amanitin (ref. 28 and H.-R.R. & H.L.S., manuscript in preparation). Assays for RNA synthesis were performed at 37 °C in a total volume of 250 µl containing Tris-HCl, *p*H 7.35 (10 mM), $MgCl_2$ (6 mM), $MnCl_2$ (1 mM), $(NH_4)_2SO_4$ (40 mM), ethylene glycol (10%), unlabelled nucleoside triphosphates (NTPs) (400 µM each), 25 µCi ^{3}H-NTPs (specific activity 52 Ci mmol^{-1}) as indicated. For DNA-directed RNA synthesis (*a*) the NTP concentration of the ^{3}H-labelled NTP was raised to 40 µM by addition of the corresponding unlabelled NTP. DNA template concentration was 25 µg ml^{-1} and RNA polymerase II concentration was 0.15 µg ml^{-1}. For RNA-directed RNA synthesis (*b*) which proved to be two orders of magnitude lower than the DNA-directed RNA synthesis when the same assay conditions were used, the corresponding ^{3}H-labelled NTP was used undiluted. RNA template concentration was 5 µg ml^{-1} and RNA polymerase II concentration was 0.75 µg ml^{-1}. Aliquots (45 µl) of the reaction mixture were withdrawn at the indicated times and spotted on Whatman 3 MM paper strips (1.5 × 10 cm) which were subsequently chromatographed in saturated $(NH_4)_2SO_4$/1 M sodium acetate, *p*H 5.2/isopropanol (50/50/2). The start was cut out and radioactivity was counted in toluene/PPO/POPOP. For each assay five samples were analysed. Circular (C PSTV) RNA and linear (L PSTV) RNA were isolated according to ref. 15.

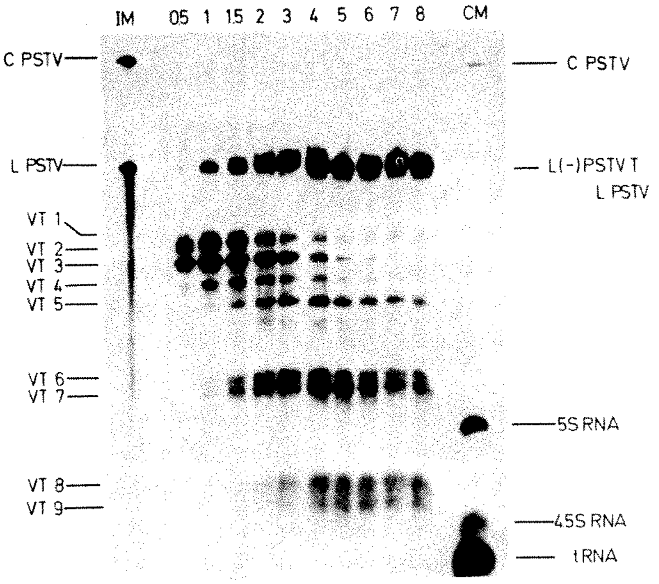


Fig. 2 PAGE analysis of the kinetics of viroid (PSTV)-directed RNA transcription by DNA-dependent RNA polymerase II from tomato. Transcription products after different times of incubation. Conditions for RNA synthesis were as described in Fig. 1 legend. The numbers on top of the slots give the time of incubation in hours. Aliquots (10 µl) were taken at the indicated time intervals and subjected to PAGE in fully denaturing conditions as described in Fig. 1*b* legend. Individual bands were visualized by fluorography of dried gels. C PSTV: ^{125}I-labelled circular PSTV; L PSTV: ^{125}I-labelled linear PSTV; L PSTV T: linear ^{32}P-labelled PSTV full-length transcript. VT1 migrates close to VT2 and due to overexposure the two bands cannot be discriminated in this autoradiogram.

Wheat-germ polymerase II seems to transcribe all three viroids practically as efficiently as the tomato enzyme (Table 1), if the overall incorporation is considered, although it is much less efficient in producing full-length viroid (PSTV) RNA copies than is tomato polymerase II (Fig. 1*b*, lanes 1, 2 and Fig. 3).

Analysis of viroid transcripts

The viroid-directed RNA ^{32}P-labelled transcription products were analysed on 5% polyacrylamide gels containing 8 M urea[1]. When the electrophoresis was carried out in high-salt TEB buffer (89 mM Tris/EDTA/borate, *p*H 8.3) and at low gel temperature (10 °C), both enzymes gave a very similar RNA product pattern consisting of several distinct bands (Fig. 1*a*, lanes 1, 2). Surprisingly, transcripts are visible which are apparently larger than the corresponding viroid template itself (Fig. 1*a*, MTV). However, no reliable conclusions about the actual size of the transcripts can be drawn from this pattern, because the gel provides denaturing conditions only for conventional RNAs, whereas viroid molecules, the corresponding transcripts and their presumptive hybrid complexes with the template remain largely undenatured. If the same products are analysed in low-salt (22 mM) TEB buffer and at high gel temperature (40 °C), where viroid RNA is fully denatured (see ref. 15), circular (C PSTV) and linear (L PSTV) viroid molecules are clearly separated (Fig. 1*b*, lanes 3–5), and it becomes evident that the largest transcripts from both polymerases have the same migrating properties and hence the same size as the linear viroid molecule. That this RNA is a genuine transcript and not due to terminal labelling of linear viroid molecules was demonstrated by the absence of transcription on omitting one of the labelled triphosphates, and by molecular hybridization.

In addition to the full-length copies about nine smaller distinct viroid transcription products are found of which only transcripts VT1–VT4 are visible in Fig. 1*b*. Although the tomato and the wheat-germ enzyme synthesize the same size classes of transcripts, densitometric analysis of the autoradiograms shows that their relative amounts differ. After 4 h of RNA synthesis 30% of the total label of the five largest RNA transcripts is found in the tomato enzyme full-length viroid copy, compared with only 9% in the wheat-germ enzyme full-length transcript.

More detailed analysis of the kinetics of tomato polymerase II transcription reveals two classes of RNA products (Fig. 2*a*): full-length viroid transcripts (L(−)PSTV T) corresponding to the linear viroid molecule (L(+)PSTV), and viroid transcription products—VT5–VT9 which accumulate with time, and transcripts VT1–VT4 (see also Fig. 1*b*) which appear very early after initiation of RNA synthesis and decrease subsequently.

Further analysis of the relationship between the different transcriptional intermediates (data not shown) revealed that their assembly into full-length copy occurred within the first 10 min after initiation. Although this rapid synthesis prevents us from determining reliably the actual sequence of the *in vitro* processing by the tomato polymerase, pulse-chase experiments and scanning of the gels clearly demonstrated that the radioactive counts are transferred from species VT1–VT4 to the full-length transcripts. As Fig. 3 shows, in the *in vitro* wheat-germ polymerase system the conversion does not proceed sequentially from the smallest transcript via intermediate-sized molecules to the full-length copies. VT2 and VT3 appear immediately after initiation followed by the formation of VT1 and VT4 (see also Fig. 1*b*) and the smaller transcripts VT5–VT9 arise at later stages of transcription. These smaller products are also found during transcription with the tomato enzyme (see Fig. 2). Although they are undoubtedly viroid specific, it is still unclear whether or not they are artefacts of the *in vitro* system.

Inhibition and binding experiments

To confirm the highly specific *in vitro* interaction between viroid RNA and polymerase II, inhibition studies with α-amanitin were carried out. This toadstool toxin specifically inhibits polymerase II at a concentration ~10^{-8} M *in vitro* and *in vivo*[16]; it has no effect on polymerase I[17], but inhibits DNA-dependent RNA polymerase III at 10^{-5} M. RNA-dependent RNA polymerase from healthy tomato tissue, also possibly involved in viroid replication, is insensitive to α-amanitin[18]. Figure 4 shows that the transcription of viroid (PSTV) RNA by the tomato polymerase II is inhibited *in vitro* by 10^{-8} M α-amanitin, providing clear evidence that the *in vitro* transcription of the viroid molecule proceeds with the aid of DNA-directed RNA polymerase II and is not based on the contaminating activity of any other known nucleic acid-synthesizing enzyme.

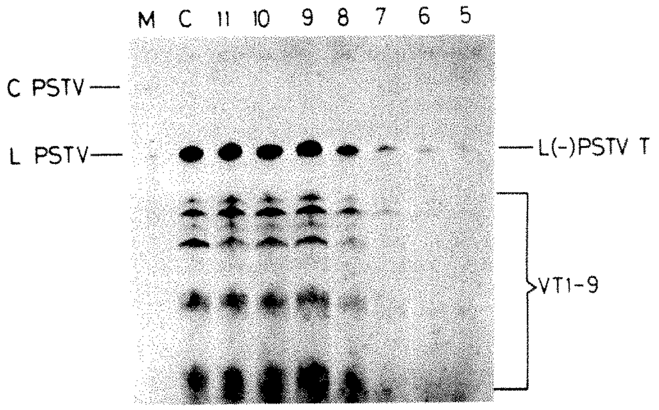


Fig. 4 Inhibition of viroid (PSTV)-directed RNA transcription by α-amanitin as analysed by PAGE. Assay conditions were as described for Fig. 1 in the absence (lane *c*) and presence of various concentrations of α-amanitin as indicated on top of the lanes, where the numbers correspond to the negative $\log_{10}$ of the molar concentrations of α-amanitin present in the assay. The assay mixture was incubated for 2 h and 10-μl aliquots were analysed on the gel as described in Fig. 1*b* legend. C, Control without α-amanitin; M, circular and linear ^{125}I-labelled PSTV marker RNA; VT1-VT9, viroid (PSTV) RNA-directed RNA transcripts.

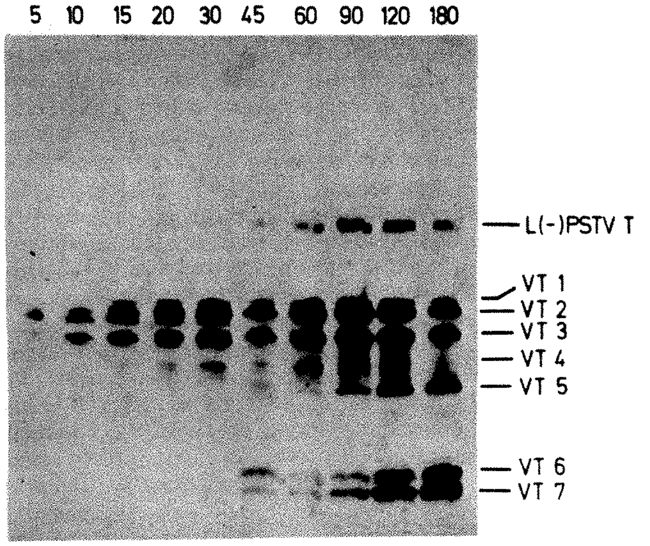


Fig. 3 Analysis of the kinetics of viroid (PSTV)-directed RNA transcription by DNA-dependent RNA polymerase II from wheat germ. Experimental conditions were as described in Figs 1 and 2 legends. The numbers on top of the slots give the time of incubation in min. The kinetics of transcription with the wheat-germ polymerase resemble a 'slow motion picture' of the more rapid processing by tomato polymerase where all transcriptional intermediates are found within 10 min after initiation. L (−) PSTV T: linear full-length viroid (PSTV) transcript, VT1–VT7: viroid (PSTV) RNA-directed RNA intermediates.

We investigated the complex formation between ^{125}I-labelled viroid (PSTV) RNA and tomato polymerase II using the nitrocellulose filter binding technique[19], in which the complex of nucleic acid template and polymerase is retained on the nitrocellulose whereas uncomplexed template is washed through the filter.

Figure 5*a* shows that a binary complex between the viroid template and the polymerase II is formed and that binding is a function of the polymerase concentration. Both circular and linear viroid (PSTV) molecules show identical binding properties. Furthermore, the binding of ^{125}I-labelled PSTV RNA to the polymerase decreased with increasing concentrations of poly-d(A-T) (Fig. 5*b*), suggesting that viroid RNA binds to the same template-binding site on the polymerase II molecule as do genuine DNA templates. This was tested with poly d(A-T) as template in the presence of viroid RNA; in this case only UTP and ATP are required for RNA synthesis and so viroid RNA-directed RNA transcription requiring all four nucleoside triphosphates will be suppressed in this assay. With increasing viroid RNA concentrations, poly r(A-U) synthesis is indeed drastically reduced (Fig. 5*c*). All these results demonstrate a specific interaction between viroid RNA and polymerase II.

Molecular hybridization

Because of the high self-complementary of the viroid molecule, experimental conditions were used which prevented self-annealing during hybridization. Thus, RNAs were separated on 5% polyacrylamide gels in fully denaturing conditions (see Fig. 1*b* legend) and after alkali-treatment were blotted and covalently linked to diazobenzyloxymethyl (DBM)-paper according to Alwine *et al.*[20]. Cellular RNAs from both healthy and PSTV-infected plants were fractionated into 2 M LiCl-soluble and -insoluble (ribosomal) RNAs and subsequently treated with DNase I to remove any DNA contamination. Radioactively labelled viroid and cellular RNAs were run as markers on the same gel.

Two approaches were taken: (1) blotting of unlabelled viroid and cellular RNAs followed by hydridization with ^{32}P-labelled PSTV transcripts, and (2) blotting of viroid, cellular and transcribed RNAs and hybridization with ^{125}I-labelled PSTV (Fig. 6*a* and *b*, respectively). From both analyses we conclude: First,

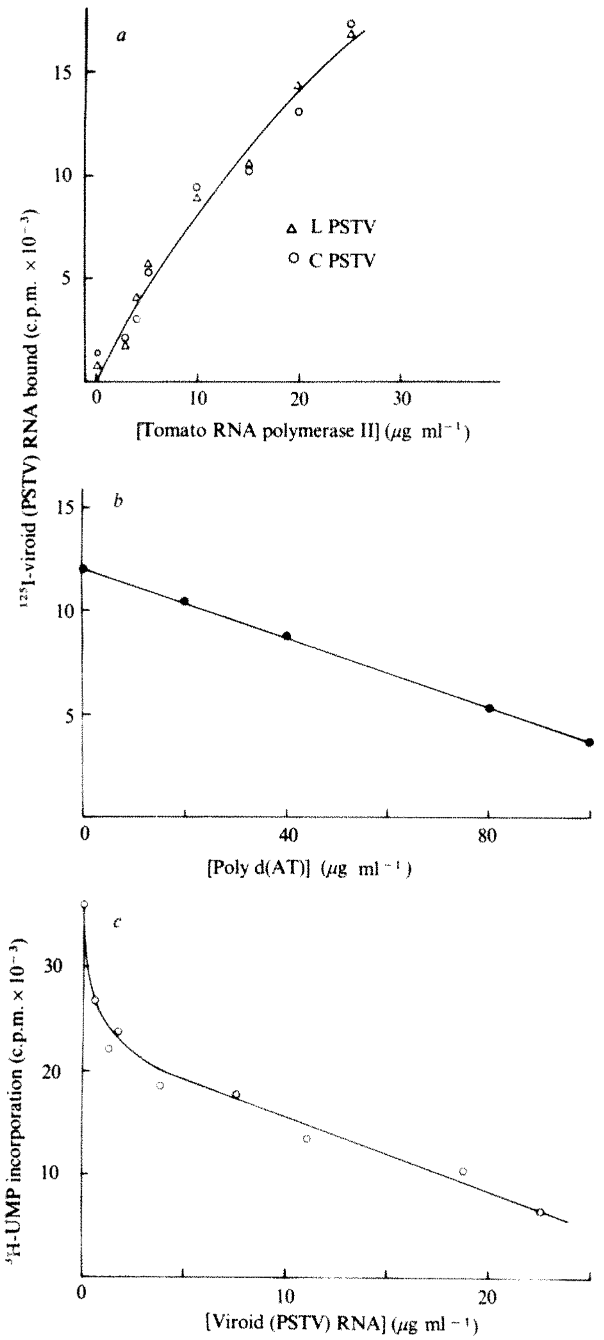


Fig. 5 Analysis of binding between DNA-dependent RNA polymerase II from tomato and viroid (PSTV) RNA. Nitrocellulose filter assay to demonstrate the binary complex formation between tomato DNA-dependent RNA polymerase II and viroid (PSTV) RNA by binding experiments (*a*) and by competition experiments (*b*) with poly d(A-T). The binding experiments were performed as described by Seidman *et al.*[19], but the polymerase assay buffer (see Table 1 legend) containing 5% dimethyl sulphoxide DMSO was used as binding buffer. The total amount of ^{125}I-labelled viroid (PSTV) RNA was 35,000 c.p.m. (specific activity 3.5×10^7 c.p.m. per μg) per individual assay. For the competition experiment the concentration of viroid (35,000 c.p.m.) and of tomato polymerase II (15 μg ml^{-1}) was kept constant whereas the poly d(A-T) template was added to the assay mixture as indicated in *b*. *c*, Inhibition of tomato DNA-dependent RNA polymerase II transcription by viroid (PSTV) RNA. The assay was performed as described in Table 1 legend. The only NTPs present in the reaction mixture were ATP (400 μM) and ^{3}H-UTP (40 μM). Polymerase concentration was 1.5 μg ml^{-1}. The assay mixture containing viroid (PSTV) RNA as indicated in *c* and all compounds necessary for RNA synthesis except DNA template was preincubated 10 min at 37 °C. The transcription was then started by the addition of poly d(A-T) template (25 μg ml^{-1}) and the assay mix was incubated for 20 min at 37 °C.

transcription of PSTV by DNA-directed RNA polymerase II yields genuine RNA transcripts which hybridize to unlabelled linear and circular PSTV (Fig. 6*a*, lane 4). This is also evident from the reciprocal hybridization of ^{125}I-PSTV to covalently fixed, unlabelled PSTV transcripts (Fig. 6*b*, lanes 4, 6). Only a few of the viroid precursor RNAs are visible (Fig. 6*b*, lanes 4, 6), because their concentration is low relative to that of the viroid template still present in the reaction mixture. Furthermore, most of the transcripts are full-length RNAs because of the increase of reaction time and ribonucleoside triphosphate concentration. Second, unlabelled linear and circular PSTV hybridize not only with their ^{32}P-labelled complementary (−)-RNA transcripts (Fig. 6*a*, lane 4), but also intermolecularly with ^{125}I-labelled PSTV exhibiting the same (+) polarity (Fig. 6*b*, lane 5), thus clearly demonstrating that molecular hybridization in standard conditions cannot distinguish between (+) and (−) viroid sequences, and therefore previously published data on this question[21,22] must be interpreted with caution. Third, whereas healthy tomato plants (Fig. 6*a*, lanes 2, 3 and 6*b*, lanes 7, 8) do not show any detectable viroid (+) or (−) sequences in the 2 M LiCl-soluble and -insoluble fractions, PSTV-infected tomato plants contain RNA species in both soluble (Fig. 6*a*, lane 5; Fig. 6*b*, lane 2) and insoluble (Fig. 6*a*, lane 6; Fig. 6*b*, lane 3) fractions that hybridize with (+) and (−) viroid RNA. These presumptive replicative intermediates occurring *in vivo* are large and do not enter the 5% gel, but can be resolved on 3.5% agarose gels into at least seven species larger than viroid RNA (see ref. 30). The exact precursor relationship, however, between these larger, viroid-specific RNA species and linear and/or circular PSTV remains to be established.

The accumulation of predominantly circular viroid molecules in plant tissue[1,15] indicates that the process of viroid replication via linear replicative (−) and (+) intermediates must involve a final step of circularization. Irrespective of this still unknown *in vivo* ligation, 5′-^{32}P-labelled linear PSTV can be ligated *in vitro* by the conventional bacteriophage T4-induced RNA ligase to form circular PSTV molecules (Fig. 6*c*).

Discussion

From the data in this and our previous papers we conclude that the first step in the replication process—the copying of the infecting viroid RNA molecule—proceeds with the aid of the DNA-dependent RNA polymerase II of the host plant. The finding that a DNA-dependent polymerase can quite efficiently accept and transcribe viroid (+) RNA into its complementary (−) RNA form may seem surprising. However, the abilities of viroid RNA to form a binary complex with the polymerase, to compete with DNA for the template binding site on the enzyme, and to inhibit DNA-directed RNA synthesis strongly suggest that the previously described DNA-like features of the viroid molecule are involved in its interaction with the polymerase. Therefore, it is intelligible from a mechanistic point of view that viroid RNA is capable of directing DNA-dependent RNA polymerase for its replication.

From the appearance of the distinct viroid-specific RNA transcripts VT1–VT4 during synthesis of the full-length viroid copy we conclude: first, the secondary structure of the viroid molecule causes the enzymes to pause at specific sites analogous to the transcription pattern of Qβ midivariant MDV-1 RNA by Qβ replicase[23], and second, specific recognition sites must exist in the viroid RNA molecule for both RNA polymerases. It is unclear why the polymerase stops after one round of *in vitro* transcription, whereas transcription products larger than the viroid molecule itself are detectable *in vivo*. Further experiments are needed to decide whether the smaller transcripts VT5–VT9 which accumulate later in transcription are caused by a site-specific nuclease activity possibly present in the polymerase preparation or whether they are transcripts of specifically nicked viroid molecules, in which case the transcription would stop at the nick.

The demonstration that viroids replicate and accumulate in the nucleus[8,24–26], and most efficiently in rapidly dividing plant tissue[27], may now be related to the fact that, due to its normal function of synthesizing genomic messenger RNAs, polymerase II is highly active in the nuclei of dividing cells, and it would

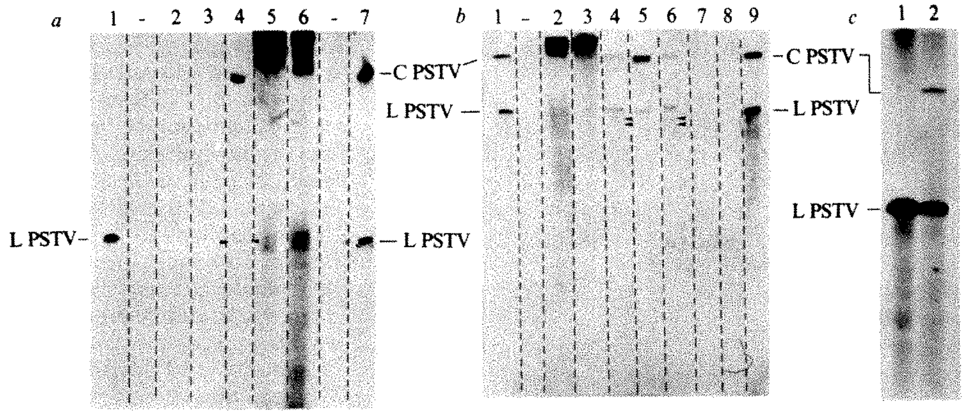


Fig. 6 Molecular hybridization of radioactively labelled (+) and (−) PSTV RNAs to RNAs covalently bound to DBM-paper. Preparation of DBM-paper, blotting in 25 mM sodium phosphate buffer, *p*H 6.8, and hybridization followed the procedure of Alwine *et al.*[17]. Separation of RNAs on 5% polyacrylamide gels in fully denaturing conditions was achieved as described in Fig. 1*b* legend. 4 μg of unlabelled PSTV RNA and 500 μg of each RNA fraction were applied; linear and mixtures of linear and circular ^{125}I-labelled PSTV served as markers and were also blotted onto the paper. Hybridization was performed at 41.5 °C for 48 h with 5×10^4 c.p.m. cm^{-2} paper surface area of ^{125}I-labelled PSTV and with 3×10^3 c.p.m. cm^{-2} paper surface area of ^{32}P-labelled PSTV transcripts. *a*, Hybridization with ^{32}P-labelled PSTV transcripts. Lanes 1, 7: ^{125}I-PSTV marker; 2, LiCl-soluble RNA from uninfected tomato; 3, rRNA from uninfected tomato; 4, unlabelled viroid (PSTV) RNA; 5, LiCl-soluble RNA from PSTV-infected tomato; 6, rRNA from PSTV-infected tomato. *b*, Hybridization with ^{125}I-labelled PSTV RNA. Lanes 1, 9, ^{125}I-PSTV marker; 2, LiCl-soluble RNA from PSTV-infected tomato; 3, rRNA from PSTV-infected tomato; 4, 6, unlabelled PSTV RNA transcripts; 5, unlabelled PSTV RNA; 7, LiCl-soluble RNA from uninfected tomato; 8, rRNA from uninfected tomato. ^{32}P-labelled RNA transcripts (specific activity 400 Ci mmol^{-1}) was prepared as described in Fig. 1 legend. For unlabelled transcription products, the ribonucleoside triphosphate concentration was raised to 400 μM each. PSTV RNA was iodinated by the method of Commerford[29] to a specific radioactivity of 3.5×10^7 c.p.m. per μg, and circular ^{125}I-labelled PSTV RNA was isolated from the reaction mixture by PAGE in fully denaturing conditions (see Fig. 1*b*). *c*, Circularization of ^{32}P-labelled linear PSTV by T4-induced RNA ligase. PSTV RNA (2 μg) containing 1–2% of linear PSTV was dephosphorylated by acid phosphatase from human prostata and labelled at its 5′-position by the action of polynucleotide kinase in the presence of [γ-^{32}P]ATP (specific activity 1,000 Ci mmol^{-1}). The ligase reaction was carried out in a total volume of 25 μl containing 50 mM Tris-HCl *p*H 7.9, 20 mM $MgCl_2$, 0.1 mM ATP, 3.3 mM dithiothreitol, and 10 units of T4-induced RNA ligase. After 30 min at 37 °C, the reaction was stopped by the addition of EDTA to 50 mM, and the reaction products were analysed as described in Fig. 1*b* legend. Before autoradiography, gels were stained with methylene blue to visualize unlabelled circular PSTV. Lane 1, 5-^{32}P-labelled linear PSTV RNA before ligation; lane 2, ^{32}P-labelled PSTV RNA after ligation.

hence be available there for intensive viroid replication. Viroid pathogenesis could also be explained in these terms. As the infecting viroid molecule directs the nuclear DNA-dependent RNA polymerase II to perform its 'selfish' replication the synthesis of genomic messenger RNAs of the host cell may become inhibited or repressed. Thus, normal differentiation would be disturbed, leading to the expression of the typical symptoms of leaf malformation and retardation of growth.

We thank Dr H.-P. Mühlbach for the continuous supply of tomato callus tissue and for helpful advice, Drs F. Boege, and D. Riesner for critical discussions, Mrs M. Bahr, M. Jung, H. Muth, K. Ramm and Mr H. Thiele for technical assistance and Mrs C. Reitz for photographic and drawing work. Our investigation was supported by the Deutsche Forschungsgemeinschaft through the Sonderforschungsbereich 47, and through a Heisenberg Fellowship given to W.R.

Received 30 September 1980; accepted 25 February 1981.

1. Sänger, H. L., Klotz, G., Riesner, D., Gross, H. J. & Kleinschmidt, A. K. *Proc. natn. Acad. Sci. U.S.A.* **73,** 3852–3856 (1976).
2. Gross, H. J. *et al. Nature* **273,** 203–208 (1978).
3. Riesner, D. *et al. J. molec. Biol.* **133,** 85–115 (1979).
4. Semancik, J. S. & Geelen, J. L. M. C. *Nature* **256,** 753–756 (1975).
5. Hadidi, A., Jones, D. M., Gillespie, D. H., Wong-Staal, F. & Diener, T. O. *Proc. natn. Acad. Sci. U.S.A.* **73,** 2453–2457 (1976).
6. Diener, T. O. & Hadidi, A. *Comprehensive Virol.* **11,** 285–335 (1977).
7. Zaitlin, M., Niblett, C. L., Dickson, E. & Goldberg, R. B. *Virology* **104,** 1–9 (1980).
8. Branch, A. D. & Dickson, E. *Virology* **104,** 10–26 (1980).
9. Mühlbach, H.-P. & Sänger, H. L. *Nature* **278,** 185–188 (1979).
10. Jendrisak, J. J. & Burgess, R. R. *Biochemistry* **14,** 4639–4645 (1975).
11. Sasaki, R., Goto, H., Arima, K. & Sasaki, Y. *Biochim. biophys. Acta* **366,** 435–442 (1974).
12. Sasaki, Y., Goto, H., Wake, T. & Sasaki, R. *Biochim. biophys. Acta* **366,** 443–453 (1974).
13. Sasaki, Y., Goto, H., Ohta, H. & Kamikubo, T. *Eur. J. Biochem.* **70,** 369–375 (1976).
14. Gross, H. J., Domdey, H. & Sänger, H. L. *Nucleic Acids Res.* **4,** 2021–2028 (1977).
15. Sänger, H. L. *et al. FEBS Lett.* **99,** 117–122 (1979).
16. Wieland, Th. & Faulstich, H. *Crit. Rev. Biochem.* **5,** 185–260 (1978).
17. Weinman, R. & Roeder, R. G. *Proc. natn. Acad. Sci. U.S.A.* **71,** 1790–1794 (1974).
18. Boege, F. & Sänger, H. L. *FEBS Lett.* **121,** 91–96 (1980).
19. Seidman, S., Surzycki, S. J., De Lorbe, W. & Gussin, G. N. *Biochemistry* **18,** 3363–3371 (1979).
20. Alwine, J. C., Kemp, D. J. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5350–5354 (1977).
21. Grill, L. K. & Semancik, J. S. *Proc. natn. Acad. Sci. U.S.A.* **75,** 896–900 (1978).
22. Grill, L. K., Negruk, V. I. & Semancik, J. S. *Virology* **107,** 24–33 (1980).
23. Mills, D. R., Dobkin, C. & Kramer, F. R. *Cell* **15,** 541–550 (1978).
24. Diener, T. O. *Virology* **43,** 75–89 (1971).
25. Sänger, H. L. *Adv. Biosci.* **8,** 103–116 (1972).
26. Takahashi, T. & Diener, T. O. *Virology* **64,** 106–114 (1975).
27. Mühlbach, H.-P. & Sänger, H. L. *Biosci. Rep.* (in the press).
28. Rackwitz, H. R., Mühlbach, H.-P. & Sänger, H. L. *Hoppe Seyler's Z. physiol. Chem.* **361,** 316 (1980).
29. Commerford, S. L. *Biochemistry* **10,** 1993–1999 (1971).
30. Rohde, W. & Sänger, H. L. *Biosci. Rep.* **1,** No. 4 (in the press).

LETTERS TO NATURE

Millimetre observations of quasar Q0420 − 388

W. A. Sherwood, G. V. Schultz & E. Kreysa

Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69,
5300 Bonn 1, FRG

The quasar, Q0420 − 388, ($Z = 3.12$, $m = 16.9$, ref. 1) was detected at a wavelength of 1 mm during part of a programme of 1-mm observations of optically selected ($m < 17.5$) quasars[2]. We observed a strongly inverted spectrum and confirm the high luminosity[3]. The observations support the explanation of the radio-quiet quasar phenomenon as being due to very compact structures which are self-absorbed even at short centimetre wavelengths.

The observations were made at the prime focus of the European Southern Observatory (ESO) 3.6-m telescope on La Silla at two epochs: 9–11 September 1979 and 1–7 July 1980. The detector for both epochs was a composite bolometer (Infrared Laboratories) with a measured noise equivalent power (NEP) of $9.6 \pm 1.0 \times 10^{-5}$ W Hz$^{-1/2}$. It was cooled to 1.2 K with superfluid helium. The sensitivity, signal-to-noise ratio of 1 in 1 h was 0.7 Jy in 1979 and 0.5 Jy in 1980. The improvement was a result of changing the filter from two capacitive meshes with a maximum transmission of 45% and half-power cuton at 650 μm to one with four capacitive meshes[4] with 75% transmission and cuton at 710 μm as measured with a Michelson interferometer. The horn exit aperture determined[5] the long wavelength cutoff at 2.1 mm.

To reduce the influence of background drift the data were taken in cycles, ABBA, where first beam A was directed on the source for 10 s then, after a 3-s delay to ensure that any electrical

0028-0836/81/220301—05$03.00

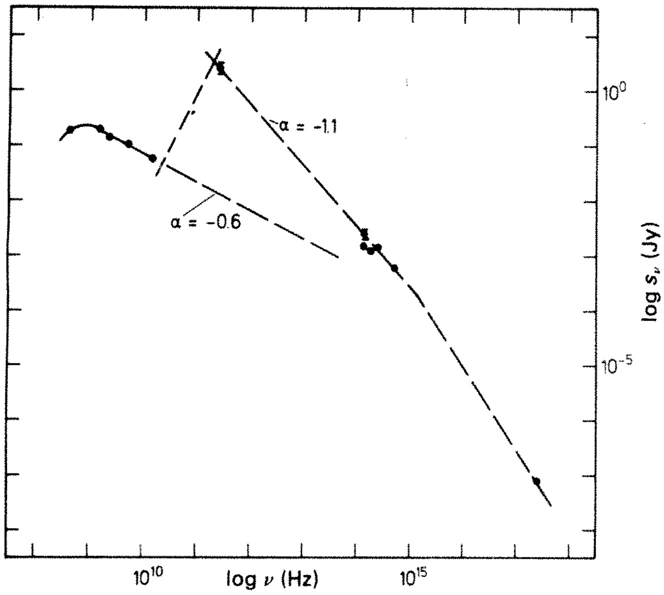


Fig. 1 The spectrum of Q0420 − 388. The radio data are from refs 15, 16 and have errors ≤10%. The IR point with the 20% error bar is from ref. 3 while the other IR data are from ref. 8 with errors ~10%. The optical[1] and X-ray[13] data also have errors of ~10%. The millimetre flux density for September 1979 is shown with an error bar of ±30%. The error in the frequency is discussed in the text.

interference had decayed, beam B was directed on source for a pair of 10-s integrations. The phase was such that, for a signal, A − B was equal to twice the true signal. At both epochs there was an excess number of positive pairs, A − B > 0, above that expected if the sign of the measurements was randomly distributed: in 1979, from 324 pairs, 186 were positive and in 1980, from 612 pairs, 401 pairs were positive. There is, therefore, no doubt that the source has been detected. The signal was chopped by means of a sector mirror mounted on an air bearing and rotating at 15 Hz (ref. 6). The measured beam size was 2 arc min (FWHP) and the separation of the beams was 2.25 arc min.

The extinction and calibration are determined from observations of the planets. The effective frequency is determined both by the spectral index of the source and by the precipitable water vapour content of the atmosphere as determined from the extinction[7]. The latter indicated 3 mm of water for the September data and the optical/IR data[1,3,8] suggested a spectral index of −1.1; the effective frequency is 245 GHz (1.2 mm). An increase of 2 mm in the water vapour content would decrease the frequency by only 6%. Increasing the spectral index to $\alpha = 0$ increases the frequency to 268 GHz and leaves the flux density unchanged. Thereafter, increasing the spectral index by 1 increases the frequency by ~8% and the flux density by ~15%.

The measured 1-mm flux densities for Q0420 − 388 were:

9–11 September 1979	2.5 ± 0.7 Jy	
1–2 July 1980	6.3 ± 0.8	(Preliminary)
3 July	3.6 ± 1.0	(Preliminary)
5 July	6.7 ± 0.9	(Preliminary)
7 July	5.2 ± 0.7	(Preliminary)

The mean values for 9–11 September 1979 and 1–7 July 1980 are significantly different and indicate that real variations have occured over ~300 days. The September value is shown in Fig. 1.

The IR data also contain the suggestion of variablility (Fig. 1). A comparison of the K magnitudes found by Wright and Kleinmann[3] (13.5 ± 0.2) and by Glass[8] (14.07 ± 0.09) show that at the 98% confidence level the magnitudes are different. Wright and Kleinmann[3] showed that during the past 40 years the optical variations have not differed from the mean by more than a factor of 2. However, photometric coverage was inadequate to detect short-term changes and the time scale for optical and IR variability remains to be determined. Our millimetre data are consistent with the size of the variations seen in the IR and optical region.

The observed radiation has a rest wavelength of ~300 μm adopting a redshift of $z = 3.12$. If this radiation were from the thermal emission of dust it would have to be a coincidence that the 1 mm flux density from the long wavelength side of the thermal peak (at ≤100 μm) agrees so well with the value extrapolated from the optical/IR region. Not only would the bolometric luminosity be considerably increased but also the luminosity of dust, in excess of $10^{15} L_\odot$, would be many orders of magnitude greater than the upper limits given for quasars or seen in galaxies or predicted for primeval galaxies even for an optimum redshift between 5 and 10 (refs 9–12). The millimetre luminosity is 10 times larger than that of the X rays[13] (Fig. 2). The addition of our data suggest that the total luminosity ought to be increased by at least a factor of 4 over that previously reported[3] to $3 \times 10^{15} L_\odot$ assuming $q_0 = 0$, and $H_0 = 50\ \mathrm{km\ s^{-1}\ Mpc^{-1}}$. ($F_{\mathrm{bol}} = (4.16 \pm 0.04) \times 10^{-14}\ \mathrm{W\ m^{-2}}$ integrating from ~24 GHz to 4.8×10^{17} Hz. Using the July data would increase these values a further 70%.)

Q0420 − 388 is not listed in the Parkes catalogue[14] but it has recently been detected between 408 MHz and 14.5 GHz (refs 15, 16), the data being obtained quasi-simultaneously over a few months in 1978. With the radio and millimetre data the quasar has an inverted spectrum between 14.5 and 245 GHz ($\alpha \geq +1.3$) (Fig. 1) and probably below ~408 MHz. This spectrum may be interpreted as originating from synchrotron self-absorbed components with turn-over frequencies near 560 MHz and 140 GHz. Although VLBI observations are required to measure the angular diameters of the components, the diameters may be estimated from the spectrum if some assumptions are made. Where VLBI data of compact components exist the magnetic fields are typically $10^{-4\pm1}$ G (ref. 17). However, in the case of Q0420-388 the magnetic energy would be insufficient to cause the synchrotron radiation. Therefore to estimate the physical size we adopt the condition that the energy in relativistic particles, E_{p}, be comparable with that in the magnetic field, E_{m}. In Table 1, the results are given, beginning with the trial value of 10^{-4} G. The magnetic field in the large component must be stronger than 10^{-3} G to achieve equipartition.

This field was used as a trial value for the small component. However, the energy in relativistic particles would then be more than 10^{11} times that in the magnetic field. A field in excess of 0.1 G is required for equality. The trend for the smallest sources to have the highest magnetic fields has been noted previously[17]. It may also be noted in Table 1 that, in this condition of approximately equal energy distributions, only 1% of the energy is in the small component (10^{55} erg compared with 10^{57} erg). Also the radiating lifetime is uncomfortably short: ~3 yr, implying a scale size 3×10^{18} cm for electrons travelling at the speed of light.

The derived linear sizes are given in Table 1 for $H_0 = 50\ \mathrm{km\ s^{-1}\ Mpc^{-1}}$ and $q_0 = 0$. For a magnetic field of 10^{-1} G the

Table 1 Physical properties of Q0420 − 388

B (G)	θ (m arc s)	E_{p} (erg)	E_{m} (erg)	$E_{\mathrm{p}}/E_{\mathrm{m}}$
Large component ($S_\nu = 0.33$ Jy at turnover $\nu = 560$ MHz)				
10^{-4}	~2.5 (1.0×10^{20} cm)	$\sim 9 \times 10^{59}$	$\sim 4 \times 10^{53}$	$\sim 2 \times 10^{6}$
10^{-3}	~4.4 (2×10^{20} cm)	$\sim 5.5 \times 10^{57}$	$\sim 2 \times 10^{56}$	~20
Small component ($S = 4.3$ Jy at turnover $\nu = 140$ GHz)				
10^{-3}	$\sim 1.6 \times 10^{-2}$ (7×10^{17} cm)	$\sim 4 \times 10^{60}$	$\sim 1 \times 10^{49}$	$\sim 4 \times 10^{11}$
10^{-1}	$\sim 5.0 \times 10^{-2}$ (2×10^{18} cm)	$\sim 1 \times 10^{56}$	$\sim 3 \times 10^{54}$	~40
(July maximum; $S_\nu = 5.8$ Jy at turnover $\nu = 300$ GHz)				
0.5	$\sim 3.4 \times 10^{-2}$ (1×10^{18} cm)	$\sim 8 \times 10^{54}$	$\sim 3 \times 10^{55}$	~0.3

Linear size based on $H_0 = 50\ \mathrm{km\ s^{-1}\ Mpc^{-1}}$ and $q_0 = 0$.

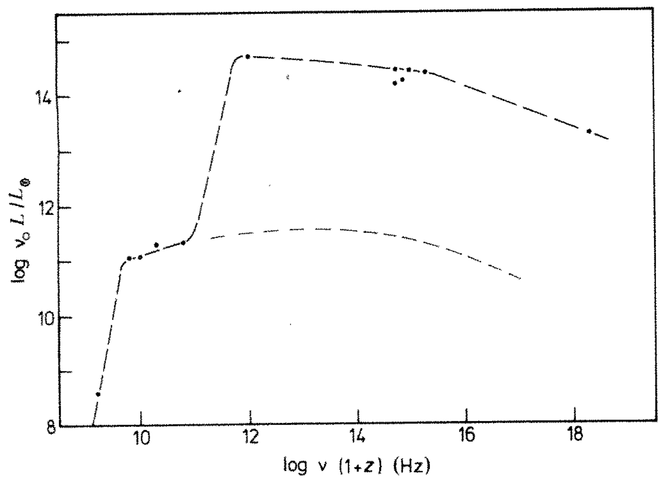


Fig. 2 Luminosity of Q0420−388. The luminosity is given by $\nu S_\nu 4\pi(c/H_0)^2 z^2(1+z/2)^2(1+z)^{-(1+\alpha)}/L_\odot$ for $H_0 = 50\ \mathrm{km\ s^{-1}\ Mpc^{-1}}$ and $q_0 = 0$, $z = 3.12$. The spectral indices were taken from Fig. 1. Note that $\alpha = 2.5$ was assumed for the 408 MHz point and $\alpha = -1.5$ for the X-ray point at 4.8×10^{17} Hz.

millimetre source has a size of 10^{18} cm in agreement with its radiating lifetime and the scale of millimetre variability. It is obvious that this point needs further study and additional confirmation.

We thank Drs Michael Arnold and Hans Peter Gemünd for work on improving filters for 1 mm, also Drs Dave Jauncey and Ian Glass for their radio and IR data and Drs Martin Harwit, Ivan Pauliny-Toth, Peter Strittmatter and Arno Witzel for discussions. This work is supported by the Deutsche Forschungsgemeinschaft, SFB 131.

Received 3 November 1980; accepted 25 March 1981.

1. Osmer, P. S. & Smith, M. G. *Astrophys. J. Lett.* **215,** L47–L50 (1977).
2. Sherwood, W. A., Kreysa, E. & Schultz, G. V. *Astr. Ges. Mitt.* **51,** (in the press).
3. Wright, E. L. & Kleinmann, D. E. *Nature* **275,** 298–300 (1978).
4. Whitcomb, S. E. & Keene, J. *Appl. Opt.* **19,** 197–198 (1980).
5. Levine, H. & Schwinger, J. *Phys. Rev.* **74,** 958–974 (1948).
6. Kreysa, E. & Schultz, G. V. *Infrared Phys.* **17,** 495–498 (1977).
7. Arnold, E. M. thesis, Univ. Bonn (1979).
8. Glass, I. S. *Mon. Not. R. astr. Soc.* **192,** 37P–39P (1980).
9. Partridge, R. B. & Peebles, P. J. E. *Astrophys. J.* **147,** 868–886 (1967).
10. Meier, D. L. *Astrophys. J.* **207,** 343–350 (1976).
11. Sunyaev, R. A., Tinsley, B. M. & Meier, D. L. *Comments Astrophys.* **7,** 183–195 (1978).
12. Koo, D. C. & Kron, R. G. *Publ. astr. Soc. Pac.* **92,** 537–545, (1980).
13. Tananbaum, H. *et al. Astrophys. J. Lett.* **234,** L9–L13 (1979).
14. Bolton, J. G. & Shimmins, A. J. *Aust. J. Phys. Astrophys. Suppl.* No. **30,** 1–54 (1973).
15. Smith, M. G. & Wright, A. E. *Mon. Not. R. astr. Soc.* **191,** 871–886 (1980).
16. Condon, J. J. *et al. Astrophys. J.* **244,** 5–11 (1981).
17. Kelllermann, K. I. & Pauliny-Toth, I. I. K. *Astrophys. J.* **155,** L71–L78 (1969); *A. Rev. Astr. Astrophys.* (in the press).

A period–luminosity relation for Mira variables in the Large Magellanic Cloud

I. S. Glass* & T. Lloyd Evans

South African Astronomical Observatory, PO Box 9,
Observatory 7935, South Africa

Mira variables are very bright objects in the IR (~2μm) so that they can be seen to large distances and through heavy interstellar absorption. Provided that they have well defined luminosities they should therefore be ideal indicators of galactic and extragalactic distances. Luminosities are also essential for understanding the physical state and pulsational characteristics of these stars which may be at the final stage of stellar evolution before planetary nebula formation[1–3]. A period–luminosity relation has been established for galactic Miras[4] but the intrinsic scatter in the relation is largely unknown. We report here periods and bolometric magnitudes for Miras in the Large Magellanic Cloud (LMC) which show a period–luminosity relation with very small scatter.

A search for red variables on visible light photographs of the Magellanic Clouds has been described elsewhere[5]; 12 Mira variables were found in a field of 0.3 deg^2 in the Bar of the LMC and periods with an error of ~10 days were determined. Details will be published elsewhere. Each star was observed at least three times in the period 1976–80 using the IR photometer[6] on the 1.9-m reflector at Sutherland. An InSb detector was used and the system is believed to be linear to better than 0.1 mag. The mean standard errors of single observations at J, H and K are 0.07, 0.04 and 0.04 mag respectively. Confusion by stars in the reference beam prevented satisfactory results for one star, which has been omitted from the discussion, and may also be responsible for the excessive red colour found for number 153. The observations do not define the light curve as a rule, so that the mean J, H and K magnitudes were found and the bolometric magnitude estimated by fitting a black-body curve[4]. The results are given in Table 1. The mean range of K in the present data is 0.5 mag, so the error in a bolometric magnitude will exceed the purely photometric errors noted above.

The 11 stars are assumed to be a homogeneous sample of M type Miras similar to those in our own Galaxy. The colours in the visible and IR regions are consistent with this except for C7 which has been found spectroscopically to be a carbon star (T. L. E., in preparation). Its exclusion would make virtually no difference to the adopted period–luminosity relation.

The data of Table 1 are plotted in Fig. 1, which also shows the best fitting straight line. This has the equation

$$m_{\mathrm{bol}} = 19.25 - 2.09 \log P \text{ (days)}$$

The r.m.s. deviation of a single star is only 0.22 mag, much of which is attributed to the small number of observations. A selection effect cannot be ruled out for the longer period stars, which are faint visually, but the sample is believed to be complete for those with $P \lesssim 300$ days. The intrinsic scatter may thus be very small and is less than the value of 0.25–0.30 mag found for classical Cepheids if no colour term is included[7]. This in turn implies that the instability strip is narrow.

The LMC distance modulus from classical Cepheids is 18.69 mag (ref. 7), so the absolute bolometric magnitudes are

$$M_{\mathrm{bol}} = 0.56 - 2.09 \log P$$

The slope of the period–luminosity relation may be fitted to the bolometric magnitude calibration for galactic Miras[4] to give

$$M_{\mathrm{bol}} = 0.76\ (\pm 0.11) - 2.09 \log P$$

Double weight has been given to the 12 stars in ref. 4 with individually known distances. The difference between the zero points of the galactic and Magellanic calibrations is not significant. Adoption of the galactic zero point would imply a distance modulus for the LMC of 18.5 mag.

The question arises whether the application of this result to distance determination would be vitiated by intrinsic differences between the stars observed in different systems. Feast[8] has argued that there is little possibility of confusion between Mira

Table 1 Bolometric magnitudes of Mira variables

Star	P (days)	m_{bol}
C38	131	14.85
C48	175	14.18
C11	200	14.40
C20	207	14.57
120	213	14.40
141	260	14.13
110	266	14.60
C7	314	14.03
153	360:	14.13:
49	390	13.67
105	430	13.51

The star names are running numbers and the numbering system has the same significance as that in ref. 5.

---

*Present address: European Southern Observatory, Karl Schwarzschild Strasse 2, D-8046 Garching bei München, FRG.

0028-0836/81/220303—02$01.00

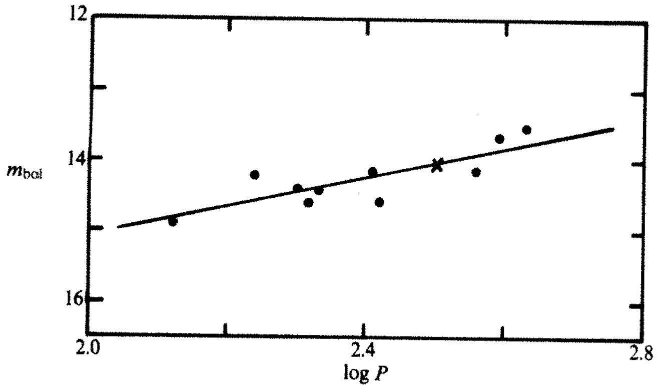


Fig. 1 m_{bol} for LMC Miras plotted against log P (days). ×, The carbon star. The best fitting linear regression line is shown.

variables and the red supergiant variables of large amplitude. There is a possibility, as in the case of other objects used as standard candles, that the luminosities depend on metal abundance. The satisfactory agreement between the zero points found for Mira variables in the Galaxy and the LMC indicates that any such difference is small.

Observations of a larger sample of stars, with more complete coverage of the light cycle, should allow an improved calibration to be found from Magellanic Cloud data alone. It seems clear from our results that Mira variables should indeed be good distance indicators.

We thank Dr M W Feast for valuable suggestions.

Received 12 January; accepted 19 March 1981.

1. Feast, M. W. in *Physical Processes in Red Giants* (eds Iben, I. & Renzini, A.) (Reidel, Dordrecht, in the press).
2. Willson, L. A. in *IAU Colloq.* No. 59 (in the press).
3. Wood, P. R. in *Physical Processes in Red Giants* (eds Iben, I. & Renzini, A.) (Reidel, Dordrecht, in the press).
4. Robertson, B. S. C. & Feast, M. W. *Mon. Not. R. astr. Soc.* (in the press).
5. Lloyd Evans, T. *The Magellanic Clouds* (ed. Muller, A. B.) 74–78 (Reidel, Dordrecht, 1971); *Mon. Not. R. astr. Soc.* **183,** 305–318 (1978).
6. Glass, I. S. *Mon. Not. R. astr. Soc.* **164,** 155–168 (1973).
7. Martin. W. L., Warren, P. R. & Feast, M. W. *Mon. Not. R. astr. Soc.* **188,** 139–157 (1979).
8. Feast, M. W. *Variability in Stars and Galaxies,* B.1.1 (Institut d'Astrophysique, Liège, 1980).

Modulation of the Earth's electric field by cosmic radiation

Ralph Markson

Center for Space Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Positive correlations between cosmic radiation and ionospheric potential (V_I) reported here indicate that the electrification of the atmosphere is modulated by changes in ionizing radiation. Analyses of two independent V_I data sets, where V_I is a measure of the intensity of the Earth's electric field[1,2], indicate that a 10% change in ground level cosmic radiation is associated with a 10–20% corresponding variation in V_I. In addition, variations in V_I were analysed relative to indices of the upper atmosphere electrical generator to investigate possible atmospheric electrical changes due to magnetospheric–ionospheric coupling; no correlation was found between the upper atmosphere indices and V_I. The nature of the global circuit is such that ionizing radiation must affect atmospheric electrification through regulation of the current output of the global thunderstorm generator which maintains the fair-weather electric field because increased conductivity in the fair-weather part of the atmosphere would lower V_I. The results also suggest the existence of a thundercloud electrification mechanism that is sensitive to ambient atmospheric electrical conditions—electric field intensity and/or conductivity[3,4].

An atmospheric electrical mechanism is a promising approach to the Sun–weather problem because it offers the possibility for solar control of energetics in the lower atmosphere without requiring energy transfer through the upper atmosphere or a change in the solar constant[5,6]. The energy distribution in the troposphere must be modified if weather is to be affected. If the Earth's electric field and/or ionization in the lower stratosphere were altered, this could influence cloud physical processes releasing latent heat and thus affecting atmospheric dynamics[7,8]. Another important characteristic of an electrical mechanism is that a rapid (<1 day) atmospheric response would be expected. Thus, it could explain meteorological effects that have been reported to follow within 1 day of solar flares[9–12] and magnetic sector crossings[13,14] which cannot be explained by heating mechanisms. The rapid response of an atmospheric electrical Sun–weather mechanism enables such a mechanism to be tested critically, whereas Sun–weather mechanisms depending on heating and planetary waves[15] have long (weeks to years) undefined time-delays and would be difficult to test experimentally.

Recently, solar wind velocity has been introduced as a parameter in Sun–weather research to examine the relationship between the fine structure of the interplanetary medium and atmospheric parameters[16]. The inverse correlation found between solar wind velocity and both cosmic radiation and V_I was taken to infer that cosmic radiation and V_I were positively correlated and that the former controlled the latter. It was postulated that discontinuities in the interplanetary magnetic field (frozen in the solar wind), which modulate cosmic radiation, cause changes in stratospheric conductivity and thus regulate currents flowing from the global thunderstorm generator to the ionosphere. These currents maintain V_I and the Earth's electric field. To investigate the assumed cosmic radiation–V_I relationship, these parameters have now been directly compared and the predicted positive correlation has been found.

Two disparate sources of cosmic radiation were correlated with ionospheric potential. The first was from the Mount Washington, New Hampshire, neutron monitor located at a height of 800 mbar (these data cover the period 1954 to the present). The median in the response curve of this instrument is 6 GeV and it is capable of sensing particles down to the geomagnetic cutoff of 0.7 GeV. The second set of cosmic ray data came from the IMP 4 satellite and constitute the longest available coherent set of satellite data including the period of the Bahamas V_I soundings used previously in the solar wind study[16]. The period June 1969–December 1972 (all the available data) was analysed. Hourly count rates were available in each data set. With the satellite data, only count rates for the most energetic particles ($E_p > 60$ MeV) were considered. Energetic particles trapped in the Earth's radiation belts and a few instances of solar protons had been removed from these data so that the record represents only the rate of galactic cosmic radiation.

The V_I data came from two independent sources. The first was obtained in the Bahamas using an aircraft during the period December 1971–March 1972[17]. Of the 94 total soundings made in the Bahamas region, 92 occurred when the Mount Washington neutron monitor data were available and 73 when the satellite was measuring cosmic radiation. Only data taken in the Bahamas and ocean regions off the coast of Florida were used in the analysis when it was found that meteorological factors inherent to conditions over land (mostly related to convection and pollution) introduced error into the 26 soundings made over the north-eastern United States, Nova Scotia and adjacent ocean areas. These factors will be discussed elsewhere.

The second V_I data set, provided by Mühleisen and Fischer[18], consists of a list of 246 V_I balloon soundings obtained mostly in Germany between 1959 and 1976. For these measurements there were 237 corresponding observations of cosmic radiation rates from Mount Washington and 79 from the satellite. The Bahamas and German V_I soundings are the only existing data sets of ionospheric potential variation that are sufficiently extensive and coherent to be useful in solar-terrestrial cor-

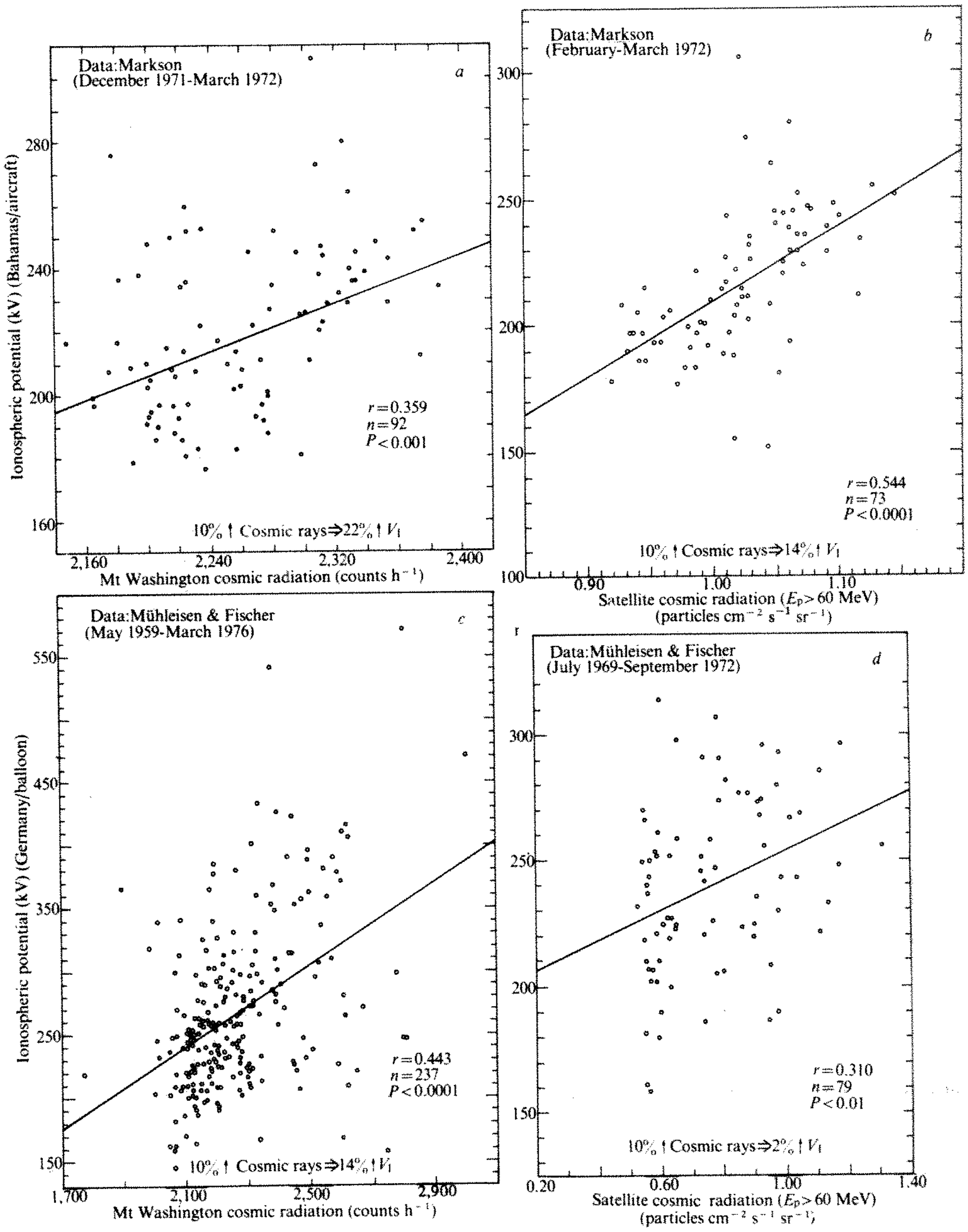


Fig. 1 Scatter diagrams of ionospheric potential (V_I) compared with cosmic radiation rates. *a*, Bahamian V_I data versus Mount Washington cosmic radiation. *b*, Bahamian V_I data versus satellite cosmic radiation. *c*, German V_I data versus Mount Washington cosmic radiation. *d*, German V_I data versus satellite cosmic radiation.

relation studies. It has been suggested that this parameter should be measured routinely or continuously to provide a "geoelectric index" to complement the "geomagnetic index"[16].

Figure 1*a* is a scatter diagram of the Bahamas ionospheric potentials compared with simultaneous Mount Washington neutron counts. The cosmic ray rates are interpolated to the times of the V_I soundings from the previous and next hour values. The correlation coefficient is 0.359 which for 92 independent data points would be significant at the 0.001 level. However, these data are not fully independent and it is difficult to estimate accurately the proper degrees of freedom for statistical testing. The effects of persistence in data on probability estimates will be considered elsewhere. Each part of Fig. 1 depicts the per cent change in the dependent variable (V_I) based on a 10% increase in the independent variable (cosmic radiation). This is determined from the slope of the least squares line of regression. In this case, a 10% increase in the Mount Washington neutron monitor rate is associated with a 22% increase in V_I.

Figure 1*b* is the complementary analysis using the satellite cosmic ray data. Here the correlation of 0.544 is larger than for the cosmic radiation measured on the ground and a 10% increase in cosmic radiation corresponds to a 14% increase in V_I.

For the balloon soundings of Mühleisen and Fischer[18], Fig. 1*c* indicates a 0.443 correlation with a 10% increase in Mount Washington cosmic radiation corresponding to a 14% increase in V_I. No interpolation was necessary with these data as their times are listed as on the hour. Figure 1*d* shows the correlation between these V_I soundings and the satellite data is 0.310 with a 10% increase in cosmic radiation associated with a 2% increase in V_I.

Figure 2*a* shows that the cosmic ray modulation was not present in the soundings made during the aircraft measuring programme in places other than the Bahamas and Florida coastal waters. These data illustrate the importance of site location in obtaining V_I measurements that are representative of global conditions.

It is of interest to compare the satellite data with Mount Washington count rates. Figure 2*b* shows that this correlation is ~0.9 with the satellite instrument twice as sensitive as the neutron monitor to changes in galactic cosmic radiation. This ratio should vary for data obtained at other phases of the sunspot cycle because of changes in the cosmic ray energy spectrum[19].

Persistence is the property of a parameter to remain unchanged for a period of time—if it is warm today it is likely to be warm tomorrow. With persistence, data in a time-series are not fully independent and this affects the degrees of freedom

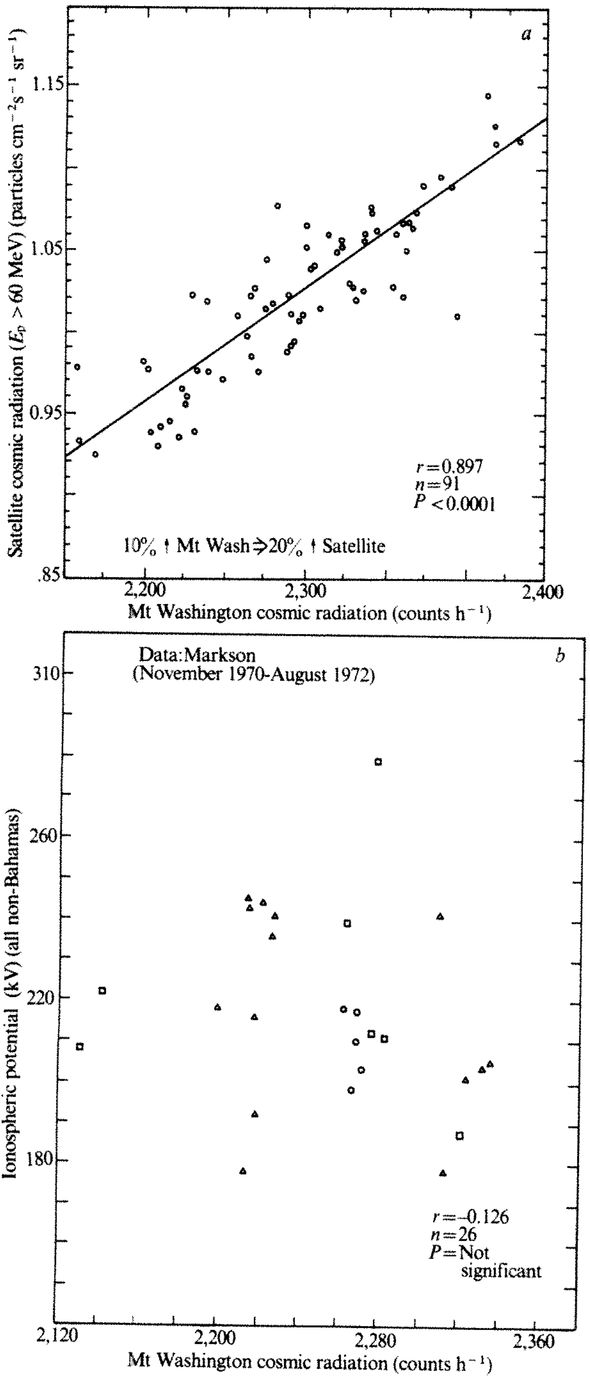


Fig. 2 *a*, Comparison of simultaneous cosmic radiation measurements obtained by the IMP 4 satellite in Earth orbit versus Mount Washington neutron monitor counts of secondary cosmic radiation at the Earth's surface. The period of these data is December 1971–March 1972. *b*, Scatter diagrams of all non-Bahamian V_I soundings obtained in the aircraft ionospheric potential investigation versus Mount Washington cosmic radiation. □, New York State and Cape Cod (November–December 1970); ○, Maryland (January 1972); ▲, Nova Scotia and New York State (April, July, August 1972).

used in computing significance. There have been too few measurements of the temporal variation of V_I to estimate accurately the persistence of this parameter, but from inspection of these records the coherency interval seems to be of the order of hours to a day. Assuming that magnetic discontinuities structured by the solar wind modulate cosmic radiation, and analysis of solar wind variations are applicable to cosmic radiation analysis, one can follow the approach of Gosling and Bame[20] who suggest that solar wind velocities a minimum of 2 days apart are reasonably independent of each other. Therefore, all the Bahamas V_I and corresponding cosmic radiation data were separated into groups that were at least 2 days apart (generally longer) and averaged. For the Bahamas measurements, Fig. 3*a* depicts a correlation of 0.486 with the Mount Washington cosmic rays and Fig. 3*b* shows a correlation of 0.663 when the satellite counts are used; these are significant at about the 0.05 level. For the larger German V_I data set, also separated into groups a minimum of 2 days apart, Fig. 3*c* shows a correlation of 0.402 with the Mount Washington data and Fig. 3*d* indicates a correlation of 0.367 for the satellite cosmic rays; the respective significance levels are 0.0001 and 0.01. Thus, analyses of two independent data sets, one obtained at relatively frequent intervals during a 3-month period in the Bahamas and the other obtained relatively infrequently over a 17-yr period in Germany, indicate that the Earth's electric field intensity is directly correlated with the rate of galactic cosmic radiation.

The importance of upper atmosphere electrical generators in modulating the Earth's electric field has been discussed elsewhere[21,22]. To investigate this the B_z component of the interplanetary magnetic field and the auroral (AE) geomagnetic index were correlated with the Bahamas ionospheric potentials. The former is highly correlated with magnetospheric–ionospheric coupled electric fields and the latter is a measure of the auroral electrojet which is maintained by the magnetospheric–ionospheric generator. Figure 4*a* is the scatter diagram for the B_z versus V_I analysis and Fig. 4*b* shows the relationship of the AE index to V_I. Both B_z and the AE index are uncorrelated with the same V_I data set that is well correlated with cosmic radiation.

Considering differences between the analyses, Fig. 1*a*, *b* for the Bahamas data shows that a 10% increase in cosmic radiation is associated with almost twice the increase in V_I as the same change in Mount Washington cosmic radiation. This may be explained by the correlation between the satellite and Mount Washington count rates (Fig. 2*a*) which exhibit a 2:1 ratio, that is, a 10% increase in neutron monitor rate would be equivalent to a 20% increase at the Earth orbiting satellite. However, comparison of Fig. 1*c* and *d* for the German data shows the magnitude of the V_I response is greater for the ground level measurements than for those at the satellite. Furthermore, comparing Fig. 1*a* and *c* the regression line slope for the German data is about half that of the Bahamas data when one correlates these V_I variations relative to the neutron monitor measurements. These discrepancies are probably due to the greatly different time span of the Bahamas data (3 months) compared with the German data (17 yr) compounded by the differences in the periods of the satellite record (3.5 yr) versus the Mount Washington record (17 yr). It is noted that during the 2-month period when the Bahamas V_I and satellite cosmic rays could be compared (Fig. 1*b*) the maximum value of cosmic ray flux was only 20% greater than its minimum value, while for the corresponding German data (Fig. 1*d*) the cosmic ray increase was 110%; the larger range of cosmic radiation variation would decrease the slope of the regression line. The different responses between the analyses is probably a function of the period length and times when the measurements were made. The cosmic ray energy spectrum is known to change in complicated ways through the sunspot cycle. This is especially true for the less energetic cosmic ray particles which are capable of ionizing the stratosphere and affect the current flow from the global thunderstorm generator to the ionosphere.

Figure 2*b* shows the critical importance of location when making V_I soundings. When sources of error associated with convection and aerosol distributions in the exchange layer were identified in the early phases of the aircraft measuring programme, it was decided to move the experimental area to the tradewind region of the Bahamas[17]. Here working in a clean maritime airmass that had travelled thousands of miles across the ocean gave measurements that were representative of global atmospheric electrical conditions.

Figure 3 shows the correlations of grouped V_I and cosmic radiation for both V_I and the two cosmic ray data sets. These groupings were done to achieve data independence. Each analysis shows that V_I and cosmic radiation are positively cor-

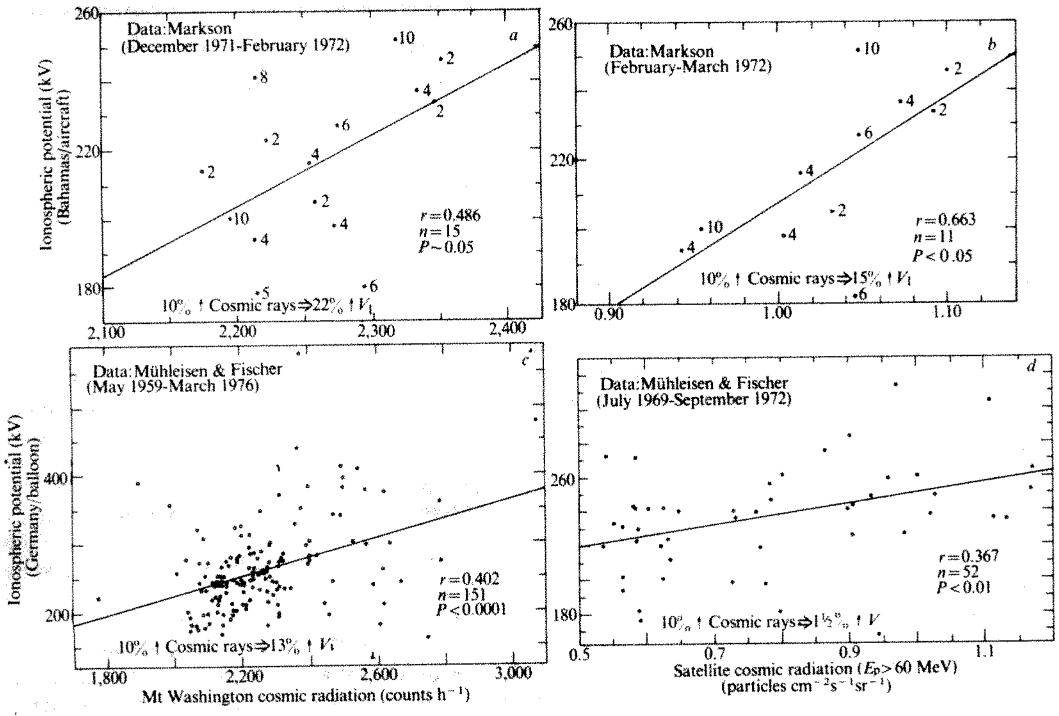


Fig. 3 Scatter diagrams of average ionospheric potentials separated into groups at least 2 days apart compared with corresponding cosmic radiation rates treated in the same manner. The numbers next to the dots give the number of data points used. *a*, Bahamian V_I data versus Mount Washington cosmic radiation. *b*, Bahamian V_I data versus satellite cosmic radiation. *c*, German V_I data versus Mount Washington cosmic radiation. *d*, German V_I data versus satellite cosmic radiation. These analyses of grouped data were performed to insure data independence.

related. The most statistically significant correlation is seen in Fig. 3*c* where the relatively large data set of Mühleisen and Fischer could be divided into groups separated by at least 2 days and still have 151 points.

The per cent increases in V_I associated with 10% increases in cosmic radiation should not be interpretated as quantitative and generalized results but rather as estimates of the range of responses that can be expected. Taken together the various correlations suggest that the ratio of ionospheric potential change to cosmic ray change is greater than unity.

Figures 4*a*, *b* which shows scatter diagrams of the relationships of B_z and the AE index to V_I, have a different appearance compared with the previous figures which showed the correlation between cosmic radiation and V_I. These parameters which are a measure of the upper atmosphere generator are uncorrelated with V_I. The correlations are not significant even though B_z and the AE index were treated as if they were independent data, which is not the case.

We conclude that solar modulation of the Earth's electric field takes place through regulation of ionizing radiation by magnetic discontinuities within the solar wind. Because galactic cosmic radiation is generally the sole agent ionizing the stratosphere and most of the troposphere (essentially all of the latter over the ocean), the Sun controls atmospheric electrification mostly

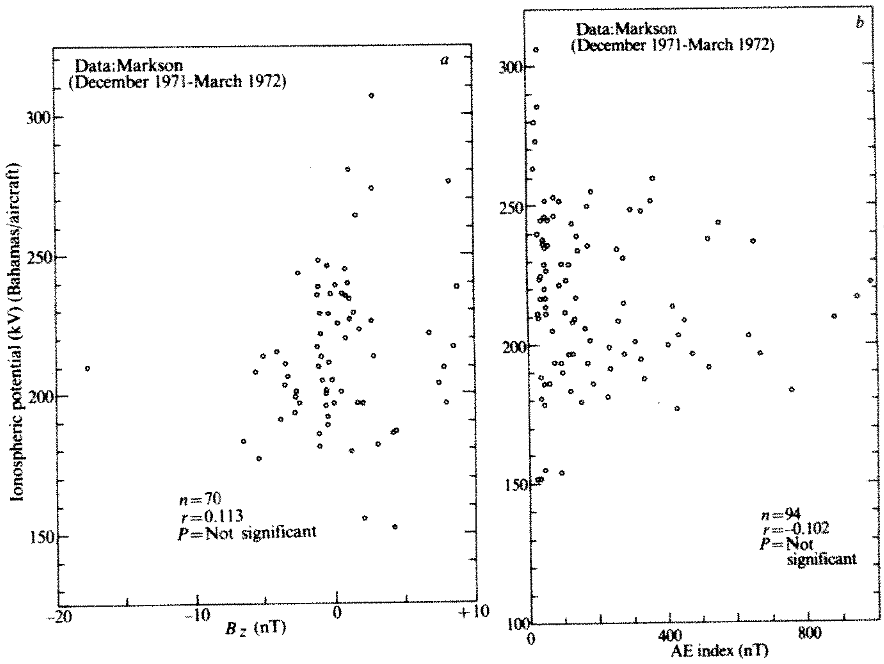


Fig. 4 Scatter diagrams of parameters representative of the magnetospheric–ionospheric electrical generator compared with V_I. *a*, Bahamian V_I data versus the B_z component of the interplanetary magnetic field. *b*, Bahamian V_I data versus the AE geomagnetic index.

through modulation of galactic cosmic radiation. However, for relatively short periods (hours to a day) after some solar flares, energetic solar particles can be the dominant ionizing source in the stratosphere[23-26]. In either case, greater ionization in the lower stratosphere would enhance thunderstorm currents and thus the fair-weather electric field. This conclusion is consistent with all past studies of the relationship of solar activity to fair-weather electric field intensity which show (1) a negative correlation between the sunspot cycle and V_I, that is a positive correlation between galactic cosmic radiation and V_I (ref 27), and (2) a positive correlation between solar flares and V_I (refs 9, 10). To examine effects due to solar activity further, I stratified the German V_I data into the years of low sunspot number (high cosmic radiation) and high sunspot number (low cosmic radiation) for comparison. The high cosmic radiation periods were 1962–66 and 1973–76 while the low cosmic radiation periods were 1959–61 and 1967–72. For the high cosmic radiation years the average V_I was 295 kV ($n = 108$); for the low cosmic radiation years the average V_I was 238 kV ($n = 142$). These data thus show a 22% peak-to-peak variation around a mean of 266 kV in phase with the solar cycle cosmic ray variation. During the cycle in question the Mount Washington neutron monitor recorded an ~17% variation around its mean.

A positive correlation between ionizing radiation and V_I implies that solar control of the global circuit takes place in the charging part of the circuit, that is, in the resistive element between the global thunderstorm generator and the ionosphere, and possibly in this generator itself. The effect of ionizing radiation on the return conduction current in the fair-weather part of the circuit would be to short out this resistive element partially and lower V_I. As increased ionization causes an increase in V_I, the effect must be due to amplification of the current output of existing thunderstorms plus possibly positive feedback in the form of increasing the number and electrification of thunderclouds. These findings support a proposed mechanism by which solar variability can modulate atmospheric electrification[6].

I thank Professor R. P. Mühleisen and Dr H.-J. Fischer for their list of ionospheric potential soundings. M. A. Shea and D. F. Smart provided the edited satellite and Mount Washington cosmic ray data. This work was supported under Office of Naval Research contract no. N00014-79-C-0399.

Received 11 December 1980; accepted 27 February 1981.

1. Chalmers, J. A. *Atmospheric Electricity* 2nd edn, 33–34, 292–296 (Pergamon, New York, 1967).
2. Israël, H. *Atmos. Electr.* **1,** 82–84 (1971); **2,** 320–324 (National Technical Information Service, Springfield, 1973).
3. Grenet, G. *Ann. Geophys.* **3,** 306 (1947).
4. Vonnegut, B. *Proc. Conf. Atmospheric Electricity* Pap. 42, 169 (AFCRL, Bedford, Massachusetts, 1955).
5. Markson, R. *Pure appl. Geophys.* **84,** 161 (1971).
6. Markson, R. *Nature* **273,** 103 (1978).
7. Battan, L. J. in *The Nature of Violent Storms*, 21 (Doubleday, New York, 1961).
8. Byers, H. R. in *Problems of Atmopsheric and Space Electricity* (ed. Coroniti, S. C.) 494 (Elsevier, New York, 1965).
9. Reiter, R. *Pure appl. Geophys.* **72,** 259 (1969); **86,** 142 (1971).
10. Cobb, W. E. *Mon. Weath. Rev.* **95,** 905 (1967).
11. Schuurmans, C. J. E. *Solar-Terrestrial Influences on Weather and Climate* (eds McCormac, B. M. & Seliga, T. A.) 105 (Reidel, Hingham, 1979).
12. Olson, R. H., Roberts, W. O. & Zerefos, C. S. *Nature* **257,** 113 (1975).
13. Wilcox, J. M., Scherer, P. H., Svalgaard, L., Roberts, W. O. & Olson, R. H. *Science* **180,** 185 (1973).
14. Larsen, M. F. & Kelley, M. C. *Geophys. Res. Lett.* **4,** 337 (1977).
15. Geller, M. A. & Alpert, J. C. *J. atmos. Sci.* **37,** 1197 (1980).
16. Markson, R. & Muir, M. *Science* **208,** 979 (1980).
17. Markson, R. *J. geophys. Res.* **81,** 1980 (1976); in *Electrical Processes in Atmospheres* (eds Dolezalek, H. & Reiter, R.) 450 (Steinkopff, Darmstadt, 1977).
18. Fischer, H.-J. & Mühleisen, R. P. *The Ionospheric Potential and the Solar Magnetic Sector Boundary Crossings* (Astronomisches Institut der Universität Tübingen, Aussenstelle Weissenau, 1980).
19. Hedgecock, P. C. *Solar Phys.* **42,** 497 (1975).
20. Gosling, J. T. & Bame, S. J. *J. geophys. Res.* **77,** 12 (1972).
21. Markson, R. *Nature* **278,** 373 (1979).
22. Roble, R. G. & Hays, P. B. *J. geophys. Res.* **84,** 7247 (1979).
23. Hake, R. D., Pierce, E. T. & Viezee, W. *Stratospheric Electricity*, 78 (Stanford Research Institute, California, 1973).
24. Winckler, J. R. *J. geophys. Res.* **65,** 1331 (1960).
25. Winckler, J. R. & Bhavsar, P. D. *J. geophys. Res.* **65,** 2537 (1960).
26. Winckler, J. R., Bhavsar, P. D. & Peterson, L. *J. geophys. Res.* **66,** 995 (1961).
27. Olson, D. E. *Symp. Influence of Solar Activity and Geomagnetic Changes on Weather and Climate* (IAGA/IAMAP Joint Assembly, Seattle, 1977).

0028–0836/81/220308—03$01.00

Neutron-capture nucleosynthesis of nature's rarest stable isotope

H. Beer

Kernforschungszentrum Karlsruhe GmbH, Institut für Angewandte Kernphysik, PO Box 3640, D–7500 Karlsruhe, FRG

Richard A. Ward

Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Strasse 1, D-8046 Garching bei München, FRG

Many attempts[1-6] have been made to explain the solar abundances of the stable, heavy odd–odd nuclei: ^{50}V, ^{138}La, ^{176}Lu and ^{180}Ta. Among these, ^{180}Ta (actually an isomeric state, $^{180}Ta^m$) has the distinction of being the rarest, stable naturally occurring isotope (0.012% of all Ta) with a solar abundance[7] of only 2.46×10^{-6} (per 10^6 Si atoms) and a half life of $>10^{13}$ yr against β decay. Although the variety of mechanisms previously proposed can in some special cases account adequately for the small abundance of $^{180}Ta^m$, it is difficult to test quantitatively those special types of synthesis invented only to explain such a small number of nuclei. In contrast, we show here from our experimental work that $^{180}Ta^m$ can also be accounted for by relatively small branchings in the much more common slow (s)- and rapid (r)-processes of stellar nucleosynthesis by neutron capture.

Because of its small abundance, the nucleus $^{180}Ta^m$ does not fit into the general systematics of s-, r- and p-process nuclei. This has been confirmed especially by recent p-process calculations[8,9] which cannot reproduce the solar abundance of this odd–odd nucleus because of the ease with which it is destroyed by photodisintegration reactions at explosive stellar temperatures. We have attempted to incorporate $^{180}Ta^m$ into the systematics of s- and r-process nuclei. Figure 1 shows nuclear structural information for the neutron-capture and β-decay processes that affect the production of $^{180}Ta^m$. Most of the information is taken from the most recent compilation in *Table of Isotopes*[10], except for the structure of ^{180}Ta which reflects recent and much-needed experimental work[11-15] on its properties. This work has not only identified many low-lying levels in ^{180}Ta but has also finally clarified which of the long-lived species in ^{180}Ta is the isomer and which is the ground state. Furthermore, the identification[13] of the 9^- nature of $^{180}Ta^m$ is very important for present considerations.

Figure 1 reveals that the key to our arguments is the population of the long-lived 8^- isomeric state $^{180}Hf^m$ (5.5 h). The possible[16] fractional β decay, $f^m_{\beta^-}$, of this level then becomes the final branch in the nuclear flow that leads directly to $^{180}Ta^m$. The initial population of $^{180}Hf^m$ in turn depends on which neutron-capture process of nucleosynthesis we are considering. If we examine the s-process neutron-capture path that flows through the valley of β stability, we see from Fig. 1 that it leads along the Hf isotopic chain through ^{179}Hf to the population of $^{180}Hf^m$. In the γ cascades down through the ^{180}Hf compound nucleus, a fraction, B, of the decays is temporarily arrested at the 5.5-h isomer. However, for the r-process, the initial flow into ^{180}Hf is somewhat different. As this process occurs far on the neutron-rich side of the valley of β stability, it is the subsequent β-decay branching at ^{180}Lu along the $A = 180$ isobaric chain that determines any r-process flow into $^{180}Hf^m$.

Considering first the possible s-process contribution, we note that the total s-process yield at $A = 180$ is usually expressed in the form of cross-section × abundance as: $\sigma N_s(A = 180) = 5.56$ ($\mathrm{mb} \times \mathrm{Si} \equiv 10^6$) according to a recent study[17] of s-process systematics at 30 keV. Therefore, the nuclear flows in Fig. 1 show that:

$$\begin{aligned} & N_s(^{180}\mathrm{Ta}^m)/N_\odot(^{180}\mathrm{Ta}^m) \\ & \quad = B\sigma N_s(A = 180) f^m_{\beta^-}/(\sigma(^{180}\mathrm{Ta}^m) N_\odot(^{180}\mathrm{Ta}^m)) \qquad (1) \\ & \quad = 2.26 \times 10^6\, B\, f^m_{\beta^-}/\sigma(^{180}\mathrm{Ta}^m) \end{aligned}$$

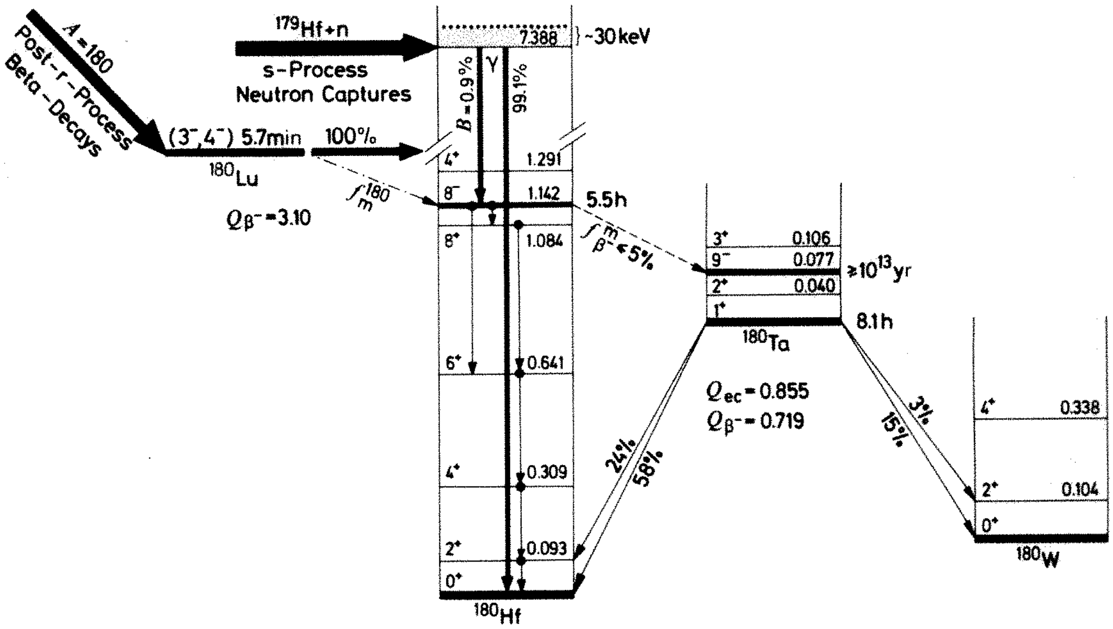


Fig. 1 The nuclear level structure along the $A = 180$ isobaric chain of significance to the stellar nucleosynthesis of ^{180}Ta^m (>10^{13} yr). The most important quantities are the fractional branchings into ^{180}Hf^m (5.5 h) through *both*: 30 keV neutron capture by ^{179}Hf (denoted by B) for the s-process and the β decay of ^{180}Lu (denoted by f_m^{180}) for the r-process. These are then followed by the possible β-decay branching, $f_{\beta^-}^m$, from ^{180}Hf^m into the isomer ^{180}Ta^m. We have also shown the spins, parities, and energies (in MeV) of many levels as well as their relevant decay properties and half lives (when known). These data are only a small representation of much recent experimental work[10–16].

for a steady flow of the neutron-capture current through $\sigma(^{180}\text{Ta}^m)$, the 30-keV neutron-capture cross-section of ^{180}Ta^m. Unfortunately, until we have an experimental determination, we must use the Hauser-Feshbach formalism[18] to calculate $\sigma(^{180}\text{Ta}^m)$. Taking into account the new information[13] that it is a 9⁻ level, we obtain $\sigma(^{180}\text{Ta}^m) = 1{,}820$ mb at 30 keV and note that this value may be uncertain by a factor of two.

As experimental input to the numerical evaluation of equation (1), we have measured the cross-section for ^{179}Hf(n, γ) ^{180}Hf^m with the activation technique. Neutrons were generated by the ^{7}Li(p, n) reaction at ~25 keV above threshold. This gives a neutron distribution in a forward cone of 120° that is very similar to a maxwellian at 25 keV thermal energy. A metallic Hf foil of natural isotopic composition back-to-back with an Au foil as a cross-section standard were both irradiated with this neutron spectrum. The induced activity was then counted using a characteristic γ-ray line with a Ge(Li)-detector. The activation foils all had a diameter of 6 mm, and the sample characteristics are summarized in Table 1. Additionally, Table 2 shows the decay properties and our measurements of the cross-sections. More details about the experimental technique can be found elsewhere[5]. The maxwellian-averaged Au cross-section was calculated from the ENDF/B-IV microscopic data file, and the measured Hf cross-section was extrapolated to $kT = 30$ keV using a theoretical calculation of Benzi *et al.*[19]. At 30 keV, the total capture cross-section of ^{179}Hf was evaluated[20] to be 1,346 mb, so we find that $B(30\ \text{keV}) = 0.90\%$, in contrast to its thermal value[21]: $B(0.025\ \text{eV}) = 0.75\%$.

With regard to possible r-process contributions, the branchings in Fig. 1 show that:

$$N_r(^{180}\text{Ta}^m)/N_\odot(^{180}\text{Ta}^m) = N_r(A = 180) f_m^{180} f_{\beta^-}^m / N_\odot(^{180}\text{Ta}^m)$$
$$= 1.22 \times 10^4 f_m^{180} f_{\beta^-}^m \qquad (2)$$

where f_m^{180} is the branching ratio of 5.7 min ^{180}Lu to ^{180}Hf^m. This branching ratio f_m^{180} could be a β-decay from ^{180}Lu directly to the 8⁻ isomer and/or β decays to higher-lying levels in ^{180}Hf followed by γ transitions to ^{180}Hf^m. For the numerical results given above, we have used $N_r(A = 180) \approx 3 \times 10^{-2}$ (Si ≡ 10^6) as being representative[17] of the average r-process yield along the $A = 180$ isobar.

Table 1 Sample characteristics

Sample	Composition Chemical	Isotopic	Investigated isotope	Thickness (atoms per barr)
Hf	Metal	Natural	^{179}Hf	8.04×10^{-5}
Au	Metal	—	^{197}Au	2.00×10^{-4}

Table 2 Decay parameters and cross-sections

Reaction	$T_{1/2}$	E_γ(keV)	I_γ(%)	σ(mb)* $kT = 25$ keV	$kT = 30$ keV
^{179}Hf(n,γ) ^{180}Hf^m	5.5 h	332	94.4±0.2	13.5±0.6	12.2±0.6
^{197}Au(n,γ) ^{198}Au	2.69 days	412	95.52±0.06	672.4±16.8	—

*Quoted uncertainties are calculated from: Au standard (2.5%); decay scheme (0.2%); Ge(Li) efficiency (1.5%); counting statistics (1%); divergence of neutron beam (1%); others (3%).

To complete the evaluation of both equations (1) and (2) we need to determine the value of $f_{\beta^-}^m$. From a theoretical point of view, the β^- energetics of the situation in Fig. 1 show that the ^{180}Hf^m → ^{180}Ta^m β decay has an endpoint energy of 210 keV and is an allowed Gamow–Teller transition because no parity change is involved. The decay would therefore be expected[22] to be characterized by an average log *ft* value in the range: $\langle \log ft \rangle = 5.7 \pm 1.1$. Using tabulated[23] f-values for such a decay, one then obtains an estimate of:

$$0.14\% \leqslant f_{\beta^-}^m \leqq 22\% \qquad (3)$$

with a value of $f_{\beta^-}^m \leqq 3.8\%$ corresponding to the limit given by Gallagher *et al.*[16] of $\log ft \geqslant 5.3$.

We can now make an approximate evaluation of equation (1) using the values: $B = 0.90\%$, $f_{\beta^-}^m \approx 3.8\%$, and $\sigma(^{180}\text{Ta}^m) = 1{,}820$ mb to obtain

$$N_s(^{180}\text{Ta}^m)/N_\odot(^{180}\text{Ta}^m) \approx 0.42 \qquad (4)$$

Furthermore, using $f_{\beta^-}^m = 3.8\%$, we see from equation (2) that a branching ratio as small as $f_m^{180} = 0.2\%$ can produce the entire solar abundance of ^{180}Ta^m through the r-process.

Considering how close to unity these estimates are for the production ratios for ^{180}Ta^m, we believe that there is ample motivation for measurements (or remeasurements) of the following important quantities needed in equations (1) and (2): (1) the β-decay branching of 5.5-h ^{180}Hf^m. For this purpose, highly enriched ^{179}Hf is needed; (2) the neutron-capture cross-section of ^{180}Ta^m at 30 keV; (3) the total cross-section of ^{179}Hf(n,γ) ^{180}Hf at 30 keV. As the cross-section is large, a sample of less than 1g is probably sufficient; (4) a firm upper limit on the fraction of the total $A = 180$ isobaric β decays that yield ^{180}Hf^m. This is a sensitive problem in nuclear chemistry and radiochemical milking; and (5) the detailed level structure of ^{180}Ta up to a few hundred keV.

Once these experimental improvements are available, the numerical evaluation of our arguments for the neutron-capture origin of ^{180}Ta^{m} will be on much firmer ground.

Received 6 February; accepted 25 March 1981.

1. Audouze, J. *Astr. Astrophys.* **8,** 436–457 (1970).
2. Mabuchi, H. & Masuda, A. *Nature* **226,** 338–339 (1970).
3. Hainebach, K. L., Schramm, D. N. & Blake, J. B. *Astrophys. J.* **205,** 920–930 (1976).
4. Harrison, T. G. *Astrophys. Lett.* **17,** 61–68 (1976); **18,** 7–12 (1976).
5. Beer, H. & Käppeler, F. *Phys. Rev.* C**21,** 534–544 (1980).
6. Beer, H., Käppeler, F., Wisshak, K. & Ward, R. A. *Astrophys. J. Suppl.* (in the press).
7. Cameron, A. G. W. In *Nuclear Astrophysics* (eds Barnes, C. A., Clayton, D. D. & Schramm, D. N.) (Cambridge University Press, 1981).
8. Audouze, J. & Truran, J. W. *Astrophys. J.* **202,** 204–213 (1975).
9. Woosley, S. E. & Howard, W. M. *Astrophys. J. Suppl.* **36,** 285–304 (1978).
10. Lederer, C. M. & Shirley, V. S. *Table of Isotopes* 7th edn (New York, Wiley, 1978).
11. Warde, E., Seltz, R., Costa, G., Magnac, D. & Gerardin, C. *J. Phys.-Lett.* **40,** L1–L3 (1979).
12. Ryves, T. B. *J. Phys.* G**6,** 763–769 (1980).
13. Burghardt, B., Harzer, R., Hoeffgen, H. J. & Meisel, G. *Phys. Lett.* **92B,** 64–66 (1980).
14. Sharma, K. S. *et al. Phys. Lett.* **91B,** 211–213 (1980).
15. Tarara, R. W., Zumbro, J. D. & Browne, C. P. *Phys. Rev.* C (submitted).
16. Gallagher, C. J. Jr Jørgensen, M. & Skilbreid, O. *Nucl. Phys.* **33,** 285–293 (1962).
17. Käppeler, F. *et al.* (in preparation).
18. Thielemann, F.-K. thesis, Tech. Hochschule Darmstadt and Max-Planck Inst., München (1980).
19. Benzi, V., D'Orazi, R., Reffo, G. & Vaccari, M. *Rep. RT/F*1(72)6 (Comituto Nazionale Energia Nucleare, 1972).
20. Drake, M. K., Sargis, D. A. & Maung, T. *Evaluated Neutron Cross Sections for Zr and Hf* EPRI-NP-250, (1976).
21. Mughabghab, S. F. & Garber, D. I. *Neutron Cross Sections* 3rd edn, Vol. I Brookhaven National Laboratory (1973).
22. Gleit, C. E., Tang, C.-W. & Coryell, C. D. *Nucl. Data Sheets* **13,** 5-5-109 (1963).
23. Gove, N. B. & Martin, M. J. *Nucl. DataTables* **10,** 205-317 (1971).

Formation of helium platelets in molybdenum

J. H. Evans

Metallurgy Division, AERE Harwell, Didcot,
Oxon OX11 ORA, UK

A. van Veen & L. M. Caspers

Interuniversity Reactor Institute, Technical University Delft,
Mekelweg 15, Delft, The Netherlands

Over the past two decades, helium introduced into metals either by ion implantation or by the (*n*, *α*) reaction has been studied over a wide range of concentrations and temperatures[1,2]. Early studies were relevant to fission reactor technology[3,4] but more recent emphasis has been on the role of helium in fusion reactor environments. One area of interest concerns the interaction of helium ions from the plasma with the first-wall surface while the influence of the helium from the (*n*, *α*) reaction on void nucleation, and on grain boundary bubble nucleation, is also under investigation[5]. One long-established property of insoluble helium and other inert gases in metals is their tendency to precipitate as bubbles. However, the inference that inert gas clusters always assume a three-dimensional form may not be justified. We present here experimental evidence showing that, at least in molybdenum, helium can initially aggregate in an unexpected planar form. Furthermore, we have found a bubble nucleation route involving the formation of several small bubbles from a single helium platelet.

At temperatures below 0.3 T_m, where T_m is the melting point, the injection of helium in molybdenum and several other metals, leads to ultrahigh density ($\sim 10^{25}$ per m^3) bubble formation[6] and bubble lattices are frequently observed[6–8]. However, bubble observations by transmission electron microscopy (TEM) can normally only be made at helium concentrations exceeding ~5 atomic %; at lower helium levels the high bubble and dislocation densities severely impede any attempts to study the early stages of bubble growth. Fortunately, a method to limit the normal high density of nucleation centres has been introduced using thermal helium desorption spectrometry (THDS)[9,10]. The THDS technique in its simplest form provides information on the binding of helium to traps in metals by first depositing energetic helium ions near to the metal surface and then examining desorption peaks during a temperature ramp. In the procedural variation of THDS which we consider here the metal is first injected with a low dose of 3 keV helium ions (typically 10^{16} ions m^{-2}) and then annealed to 1,050 K. This procedure results in a known concentration of HeV defects (monovacancies containing one helium atom) which then act as unsaturable traps for a further dose of low energy non-damaging 150 eV helium ions. In this way, no additional nucleation centres are formed but each initial HeV trap can acquire a controlled number of helium atoms up to several thousand. The unsaturable nature of the traps stems from a continuation of the initial trap mutation process[11] in which the combined strain of several helium atoms at a monovacant site can result in the removal of a lattice atom to an interstitial position leaving the helium atoms in a divacancy. Thus, as further helium is captured the trap size will expand appropriately. We have taken advantage of the control on the size and density of helium clusters inherent in the above approach, to prepare THDS samples suitable for subsequent TEM examination.

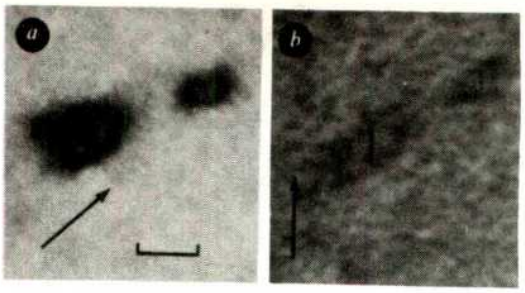


Fig. 1 Transmission electron micrographs with a beam direction close to [001] showing the edge-on projection of two helium platelets lying on (110) planes. The direction of the operating reflection, **g**, is denoted by the arrows. *a*, **g** = [020]; the platelet and loop images are visible; *b*, **g** = [$\bar{1}$10]; the loop image is out of contrast. Scale bar, 15 nm.

So far, five specimens have been prepared by this technique with a variety of initial trapping sites and different degrees of helium filling from 400 up to 4,000 helium atoms per trap. The specimens were in the form of 3-mm disks of single crystal zone-refined molybdenum with surface normals close to ⟨001⟩. A back-thinning electropolishing technique was used to obtain suitable electron thin regions in the surface layers of interest. On examination by TEM, using a Philips EM400T microscope at 120 keV, all specimens contained strain contrast loop images. In the thinner regions of foils (20–60 nm) the areal density of images was independent of thickness and matched the variation of initial trapping site density. Also the size of these strain images reflected the degree of helium trapping. Thus, they could be confidently identified with the helium traps detected by THDS. In the thicker regions of the foil, there was an increase in areal image density due to loops not associated with the traps. We anticipated the helium trap images would be due to strain fields around small over-pressurized helium bubbles; instead it became clear in three specimens that the loop images were around helium aggregates having an unexpected platelet character. The platelets lay on {110} planes and **g** · **b** = 0 criteria (where **g** is the operating reflection and **b** is the Burgers vector) showed the loops to have Burgers vectors orthogonal to their platelets. The character of the platelets was most easily seen when they were viewed close to edge-on; two such platelets are shown in Fig. 1*a* and *b* where the beam direction was close to [001]. In Fig. 1*a* with **g** = [020], the platelets are seen as a single dark line surrounded by the loop image while in Fig. 1*b* with **g** = [$\bar{1}$10] the latter image is removed (**g** · **b** = 0) to leave only the platelets visible. When viewed with the platelets inclined to the beam, the images become more complex with an inner image associated with the platelet together with the outer loop contrast. Several such images are seen in Fig. 2.

The observed platelets can be described as disk-shaped cavities containing helium. From the number of helium atoms per trap, the thickness of the platelet images ($\sim$1 nm) and their diameter, we can estimate that there are between two and three helium atoms per vacancy in the cavity. The atomic densities of helium thus seem to be $>10^{29}$ atoms m^{-3}. Similar high values have been estimated by Donnelly *et al.*[12] during optical studies of helium bubbles in aluminium. The gas pressures in the platelets are clearly immense and give rise to strain fields, and contrast effects, similar to those from intrinsic interstitial dislocation loops. The platelets can be considered analogous to the disks of helium seen in boron carbide after neutron irradiation[13–15], although the mechanics of nucleation could well be different in the different host lattice structures. More relevant may be the report of diffraction pattern streaking, attributed to argon-vacancy platelets, found in copper after bombardment with argon ions[16].

The effect of temperature on the platelets was followed by *in situ* microscope annealing. Between 600 and 800 °C the platelets transformed into helium bubbles (Fig. 3). Again, an unforeseen result emerged; the gas from a platelet did not minimize its surface energy by forming a single bubble but instead formed a cluster of small bubbles. We have mentioned that platelet formation was found only in three specimens; in the other two specimens, one with a very high initial density of helium traps and the other where each trap contained $\sim$4,000 helium atoms, separated clusters of small bubbles were found, even without annealing. The similarity between the bubble clusters in these two specimens and those in the annealed samples was close enough to be fairly certain that they too had evolved from platelets. The mechanism for the platelet to bubble cluster transformation at room temperature is uncertain. It is possible that the interaction of lattice strains from different platelets could play a part but further studies are required to establish the critical trap density and helium filling parameters.

It is interesting that the size of individual bubbles and their separation within a cluster ($\sim$3 to 5 nm) are close to those eventually found after conventional implantation to high helium concentrations. We may have uncovered a route that leads to the ultrahigh bubble density, and eventually to the bubble lattice, found in molybdenum and other metals.

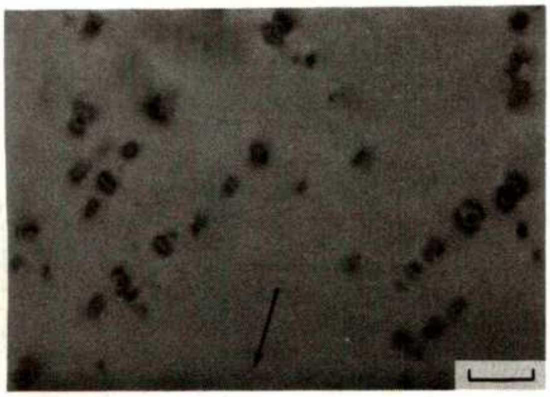


Fig. 2 Micrograph of inclined platelet/loop images using $\mathbf{g}=[\bar{1}\bar{2}\bar{1}]$ (arrowed) and a beam direction close to $[\bar{1}\bar{1}3]$. Several complex images can be seen; the simpler loop images arise from the platelet growth process. Scale bar, 40 nm.

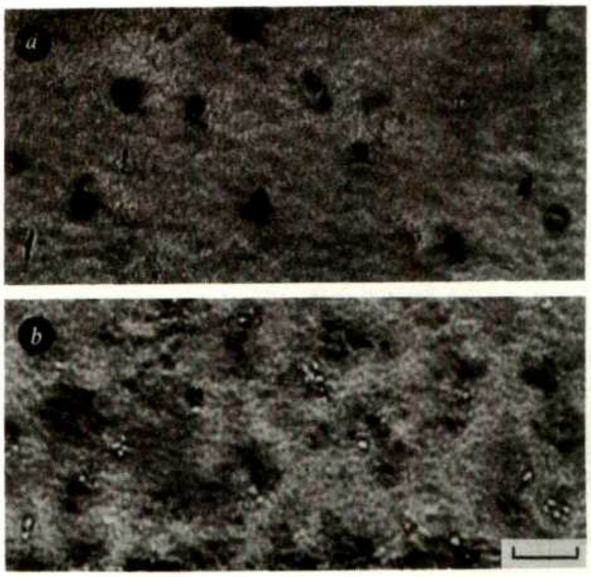


Fig. 3 Effect of annealing on platelet morphology. *a*, Unannealed; *b*, same area after 10 min at 800 °C. The platelets, some edge-on, and some inclined with loop contrast, have transformed to clusters of bubbles. Scale bar, 25 nm.

Regarding the initial platelet nucleation, by using atomistic calculations with interatomic potentials for α-iron, Caspers *et al.*[17] have demonstrated that two-dimensional clustering of helium-vacancy complexes on {011} planes is not unexpected. The results show that the self-interstitials removed from a complex by the trap mutation process do not migrate away to other sinks but remain bound to the complex and thus could have an important role in its morphology. Intermediate stages of platelet growth may not be easy to model but in later stages we have experimental evidence, described in more detail elsewhere[18], which demonstrates that the helium platelets can increase their thickness by cooperatively moving out a plane of host atoms in a glide direction. This is essentially the loop punching process proposed by Greenwood *et al.*[19] for the relief of pressure in gas bubbles. The process gives the bubble, or in our case the platelet, an extra plane of vacancies in exchange for an interstitial loop in the lattice. The characteristic of such loops is their alignment along a glide direction (⟨111⟩ in body-centred cubic metals). We found numerous instances of two to four loops aligned along such directions and having ⟨111⟩ Burgers vectors. The presence of these loops was the reason mentioned earlier for the increase in image density in thicker foil regions. Note that with our [001] specimen orientation all the ⟨111⟩ glide directions bisected the specimen surfaces. Loop loss to the surface was therefore quite likely and was sometimes observed during microscope observation. To study the loop punching process in more detail, and to establish the platelet size for its initiation, specimens with ⟨011⟩ normals would be preferred because these have two of the four ⟨111⟩ directions lying parallel to the specimen surface.

We have uncovered new and unexpected aspects of helium clustering behaviour in molybdenum and demonstrated the effectiveness of combining the THDS and TEM techniques. Some areas requiring further experiments have already been discussed such as the platelet to bubble cluster transformation and the initiation of loop punching—but the work also raises other questions. These include the effect of temperature during helium implantation and whether the platelet formation might be found with other inert gases or in other metals. However, the main interest might be to establish if the platelet formation and the possible route to bubble nucleation still operates when the deposition of helium is accomplished by the displacement damage (and the consequent high density of vacancy traps) characteristic of conventional helium implantation experiments.

We thank K. T. Westerduin for technical assistance. J.H.E. thanks the Physical and Metallurgical Departments of the Technical University, Delft for their hospitality during the initiation of this work.

Received 2 February; accepted 23 March 1981.

1. Blackburn, R. *Metall. Rev.* **11,** 159–176 (1966).
2. Pugh, S. F. (ed.) *Radiat. Effects* **53,** 105–278 (1981); *AERE Rep.* 9733 (Harwell, 1980).
3. Barnes, R. S. *Nature* **206,** 1307–1310 (1965).
4. Harries, D. R. *J. Br. nucl. Soc.* **5,** 74–87 (1966).
5. Wiffen, F. W., DeVan, J. H. & Stiegler, J. O. (eds) *J. nucl. Mater.* **85/86** (1979).
6. Mazey, D. J., Eyre, B. L., Evans, J. H., Erents., S. K. & McCracken, G. M. *J. nucl. Mater.* **64,** 145–156 (1977).
7. Johnson, P. B. & Mazey, D. J. *Nature* **276,** 595–596 (1978).
8. Johnson, P. B. & Mazey, D. J. *J. nucl. Mater.* **93/94** 721–727 (1980).
9. van Veen, A. & Caspers, L. M. *AERE Rep.* 9733, 494–517 (Harwell, 1980).
10. Kornelsen, E. & van Gorkum, A. A. *J. nucl. Mater.* **92,** 79–88 (1980).
11. Caspers, L. M., Fastenau, R. H. J., van Veen, A. & van Heugten, W. F. W. M. *Phys. Status Solidi* A**46,** 541–546 (1978).
12. Donnelly, S. E., Rife, J. C., Gilles, J. M. & Lucas, A. A. *J. nucl. Mater.* **93/94,** 767–772 (1980).
13. Jostens, A. & DuBose, C. K. H. *J. nucl. Mater.* **44,** 91–95 (1972).
14. Ashbee, K. H. G., Frank, F. C. & DuBose, C. K. H. *J. nucl. Mater.* **48,** 193–198 (1973).
15. Hollenberg, G. W., Mastel, B. & Basmajian, J. A. *J. Am. ceram. Soc.* **63,** 376–380 (1980).
16. deHosson, J. Th. M. *Phys. Status Solidi* A**40,** 293–301 (1977).
17. Caspers, L. M., Ypma, M. R., van Veen, A. & van der Kolk, G. J. *Phys. Status Solidi* (in the press).
18. Evans, J. H., van Veen, A. & Caspers, L. M. *Scr. Metal.* **15,** 323–326 (1981).
19. Greenwood, G. W., Foreman, A. J. E. & Rimmer, D. E. *J. nucl. Mater.* **1,** 305–324 (1959).

Atmospheric transport of continentally derived lipids to the tropical North Pacific

Robert B. Gagosian, Edward T. Peltzer & Oliver C. Zafiriou

Department of Chemistry, Woods Hole Oceanographic Institution,
Woods Hole, Massachusetts 02543, USA

Lipid class compounds have been used as source markers to ascertain the terrestrial, marine, or anthropogenic origin of atmospheric aerosols from rural continental and coastal oceanic environments[1–11]. However, there are only limited data on the composition and source of lipids in remote marine aerosols. Studies off the West African coast suggest that surface waxes of terrestrial vascular plants are a major source of lipid class compounds in dust samples[6,11]. Uniform concentrations of lipids found[7,10] in aerosols from urban, rural and oceanic regions imply either a long residence time for continental material or a uniform production rate of the compounds over the oceans. It has also been concluded[1–3,8] that oceanic aerosols have significant marine sources. However, previous organic chemical studies have not been conducted in conjunction with other chemical analyses of the aerosols or with climatological and meteorological studies. We report here, as part of the Sea/Air Exchange Program (SEAREX), the temporal variation of the transport of continentally-derived lipids (hydrocarbons, fatty alcohols and fatty acids) to the central tropical North Pacific.

Samples were collected during the SEAREX wet and dry season experiments (April–May and July–August 1979) from a 20-m tower erected on the windward shore of Bokandretok Island, Enewetak Atoll, Marshall Islands. The site, 11°20′ N, 162°20′ E, is located in the north-east tradewind regime, ~5,000 km south-east of Asia and 8,000 km west-southwest of the North American continent. Atmospheric particulate matter samples collected using 20×25 cm Gelman A/E glass fibre filters were pre-extracted once with acetone/MeOH and twice with CH_2Cl_2 and stored in precleaned glass jars. These filters have a collection efficiency of ≥98% for particles with radii >0.015 μm at the 20–60 $m^3 h^{-1}$ flow rate used[12]. Filters were mounted in aluminium holders within aluminium rain shelters. To prevent back diffusion or injection of organic material onto the filters, a stainless steel check valve was located between the filter frame and more than 20 feet of aluminium tubing connecting the high volume sampler to the high volume pump. Polyurethane foam plugs placed behind the filter also prevented back diffusion of organic material[12]. Quadruplicate filters were used for most sampling periods, two filters as blanks (one in each of two shelters), and two filters for duplicate samples in these separate shelters. Samples were collected for 2–10 days (2,000–10,000 m^3 air sampled). Comparison between high and low loadings showed no statistically significant difference in hydrocarbon (nC_{21}–nC_{36}), fatty alcohol (nC_{13}–nC_{32}) or fatty acid (nC_{13}–nC_{32}) concentrations. Less than 5% of the material collected on the first of back-to-back mounted filters appeared on the second filter. There were no qualitative lipid composition changes between these filters. The collection system was electronically controlled; relevant meteorological parameters were monitored continuously. Collection was stopped automatically when the surface wind direction was not from the open sea (040°–160°), when the wind speed dropped below 5 knots, when the condensation nuclei count exceeded 750 particle cm^{-3} or when precipitation started. Our sampling system has apparently been free of contamination from anthropogenic sources on the nearby main island of Enewetak, as no single polynuclear aromatic hydrocarbon (PAH) was found at concentrations

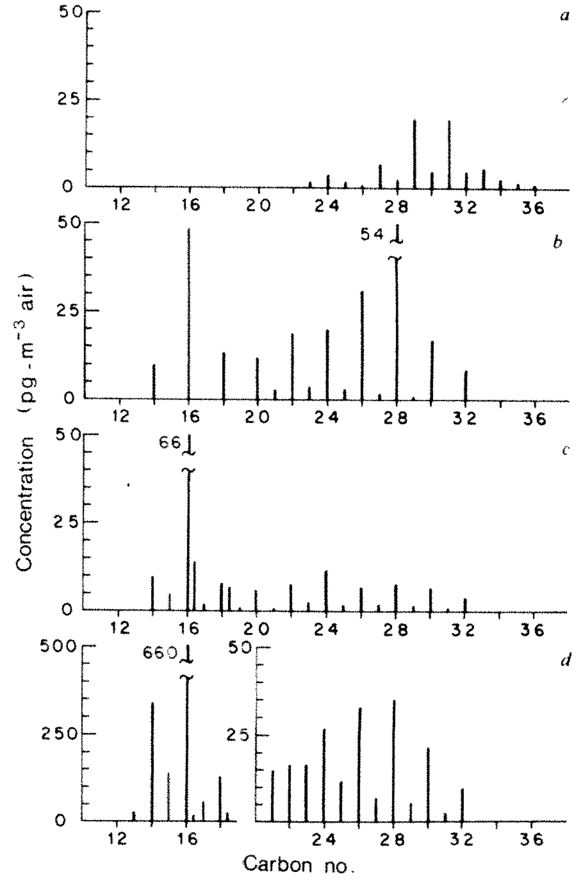


Fig. 1 Concentrations of individual aliphatic hydrocarbons (*a*), fatty alcohols (*b*), fatty acid esters (*c*) and fatty acid salts (*d*) as a function of carbon number in a marine aerosol sample collected at Enewetak Atoll, 17 May 1979. Sample volume was 4,709 m^3 collected at 75 $m^3 h^{-1}$. Peaks at half carbon numbers represent monounsaturated compounds. Detection limits for the lipid classes are 2 pg m^{-3} air in a 1,000 m^3 air sample. Blanks were run for each sample as described in the text. These values were subtracted from the samples. For the alkanes, the blank levels were 5–10 times lower in concentration (~5 pg m^{-3} for 5,000 m^3 sample) than the sample and show no odd/even carbon predominance. The fatty alcohol blanks for compounds $>C_{21}$ were <1 pg m^{-3} per component. Blanks for compounds C_{12}–C_{21} were less than five times the sample concentration (5 pg m^{-3} for a 5,000 m^3 sample) except for $C_{16:0}$ which was ~15 pg m^{-3}. Fatty acid ester blanks were <1 pg m^{-3} for compounds $>C_{21}$. $C_{14:0}$, $C_{16:0}$, $C_{18:0}$ were less than 15 pg m^{-3}. Fatty acid salt blanks averaged <5 pg m^{-3} for components $>C_{21}$. $C_{14:0}$, $C_{16:0}$, $C_{18:0}$ had blanks as high as 100 pg m^{-3} but were ~3 times lower in concentration than found in the samples.

0028-0836/81/220312—03$01.00

Table 1 Atmospheric aerosol concentrations of aliphatic hydrocarbons, fatty alcohols, fatty acid esters and fatty acid salts from samples collected at Enewetak Atoll, April–August 1979

Mid-date	18 April	6 May	17 May	5 July	14 July	29 July
Volume (m^3)	5,300	4,800	4,700	4,500	8,500	9,200
Aliphatic hydrocarbons						
C_{21}–C_{36}*	160	63	79	72	23	20
CPI‡ (C_{21}–C_{36})	3.2	2.2	2.8	3.8	2.5	4.0
Major components†	29, 31, 27	31, 29, 33	29, 31, 27	29, 31, 27	29, 31, 33	31, 29, 27
Fatty Alcohols						
C_{13}–C_{20}*	4	2	85	—	25	12
Major components†	20, 16	20, 14	16, 18, 20	—	16, 18, 14	16, 14, 20
C_{21}–C_{32}*	170	72	160	210	58	60
Major components†	28, 26, 30	28, 30, 32	28, 26, 24	30, 28, 32	30, 28, 26	30, 28, 32
CPI‡ (C_{21}–C_{32})	8.6	12	12	15	21	14
Fatty acid esters						
C_{13}–C_{20}*		290	120	90	59	34
Major components†		16, 18, 14	16, 16:1, 14	16, 14, 18	16, 14, 18	16, 14, 15
C_{21}–C_{32}*		91	56	75	10	6
Major components†		24, 22, 26	24, 22, 28	24, 22, 26	24, 22, 26	24, 20, 22
CPI‡ (C_{21}–C_{32})		6.4	4.5	7.1	7.3	12
Fatty acid salts						
C_{13}–C_{20}*	3,200	4,000	1,400	87	730	1,400
Major components†	16, 14, 15	16, 14, 18	16, 14, 15	14, 18:1, 16:1	16, 14, 18	16, 14, 15
C_{21}–C_{32}*	670	250	200	140	36	53
Major components†	28, 22, 30	28, 24, 30	28, 26, 24	28, 30, 26	30, 22, 24	26, 22, 24
CPI‡ (C_{21}–C_{32})	2.6	1.8	2.4	4.0	4.1	>10

*Concentrations are in pg m^{-3} air.

†Carbon number in order of decreasing abundance.

‡CPI, carbon preference index. Odd/even for aliphatic hydrocarbons, even/odd for fatty alcohols and acids.

>5 pg m^{-3} (see ref. 9 for an example of PAH aerosol values from urban, rural and seashore areas). No insects were detected on the filters and no branched hydrocarbons (>C_{25}), indicative of insect cuticles[13], were detected on our filters (Fig. 1), suggesting that contamination from insects trapped on the filter was minimal.

After sample collection, the filters were spiked with hydrocarbon, fatty alcohol and fatty acid internal standards for recovery determinations, and were solvent extracted with 400 ml of glass-distilled CH_2Cl_2 for 16–20 h. (Extraction of the filter with another 400 ml portion of CH_2Cl_2 yielded <1% of the hydrocarbons and fatty alcohols found from the first extraction.) The CH_2Cl_2 extract was evaporated to dryness under vacuum at <25 °C. An aliquot of the extract was separated into individual compound classes by adsorption chromatography by elution with a series of solvents of increasing polarity (ranging from hexane to 50/50 hexane/toluene to 20% ethyl acetate in hexane) to separate the lipids into alkane and alkene, polynuclear aromatic hydrocarbon, fatty acid ester, fatty alcohol, and steroid alcohol fractions. The alcohols were converted to their acetates with acetic anhydride in pyridine. The individual fractions were spiked with appropriate external standards for quantitative determinations and analysed by high resolution glass capillary gas chromatography (GC) using a deactivated (by persilylation) SE-52 coated (20.2 m × 0.32 mm i.d.) column[14]. Compound structures were identified by comparison of GC retention times and mass spectra with authentic standards.

Another aliquot of the CH_2Cl_2 extract was treated with BF_3/MeOH to convert fatty acid esters (wax esters and triglycerols) into their methyl esters (free fatty acids remained on the filter as salts). The transesterified aliquot was then treated as described above. After methylene chloride extraction, the filter was treated with 0.1 M HCl/MeOH to protonate all weak acid salts. The acidic MeOH was then extracted with hexane and the liberated fatty acids were converted to their corresponding methyl esters and analysed as described above.

Aliphatic hydrocarbons, fatty alcohols, fatty acid ester and fatty acid salt concentrations in pg m^{-3} are listed in Table 1 for six sampling periods throughout the five-month Enewetak experiment. Concentrations of the individual members of the lipid classes and their compound distributions are shown in Fig. 1 for one sample taken at the end of the dry season. Total aliphatic hydrocarbon (C_{21}–C_{36}) concentrations range from 20 to 160 pg m^3 (~3% of the total CH_2Cl_2 extract). An odd-to-even carbon preference index of 2.2–4.0 and the major alkanes being nC_{27}, nC_{29} and nC_{31} strongly suggest vascular plant wax alkane inputs[15,16]. Only traces of aliphatic hydrocarbons <C_{20} were detected, possibly due to loss from the filter of these more volatile compounds. No indication of any 'hump' from fossil fuel material was observed in the gas chromatograms of the hydrocarbon fractions[12,17]. Very low levels (<5 pg m^{-3}) of >C_{20} n-alkenes or branched alkanes were found. The increased reactivity of alkenes relative to alkanes towards photolytic or chemical oxidation may account for these low levels.

The total fatty alcohol (C_{13}–C_{32}) concentrations range from 72 to 245 pg m^{-3} (~9% of the total CH_2Cl_2 extract). The very strong even-to-odd carbon preference index of 8.6–21 for fatty alcohols >C_{19} and the presence of major amounts of nC_{26}, nC_{28} and nC_{30} compounds again suggest higher plant waxes as a source of the aerosol lipids[16]. Only a few peaks of low concentration were attributable to unsaturated fatty alcohols or fatty alcohols derived from marine sources such as $C_{12:0}$, $C_{16:0}$, $C_{16:1}$ and $C_{18:1}$ (in $C_{x:y}$ x = number of carbons and y = number of double bonds). No major (<20%) difference was observed in fatty alcohol concentration when the BF_3/MeOH treated CH_2Cl_2 extract (before column chromatography) was compared with the untreated extract. Hence, few n-alcohols were present in the form of wax esters. As BF_3/MeOH is known to destroy some unsaturated fatty alcohols such as phytol, we cannot rule out the possibility that unsaturated n-alcohols were esterified as waxes. Low concentrations (<10 pg m^{-3}) of steroid alcohols were detected; cholesterol was the major component.

Total esters (wax esters, triglycerols, and so on) of C_{13}–C_{32} fatty acids range from 40 to 381 pg m^{-3} (~7% of the total CH_2Cl_2 extract), whereas the fatty acid salts have a range of 227–4,250 pg m^{-3} (>50% of the total acidic MeOH/hexane extract). Both fractions exhibit high even-to-odd carbon preference indices for C_{21}–C_{32}. Major components for both fatty acid esters and salts >C_{19} are $nC_{22:0}$, $nC_{24:0}$, $nC_{26:0}$ and $nC_{28:0}$ again indicative of a natural terrestrial source[16]. However, the fatty acid fractions also have high concentrations of lower carbon number homologues (C_{13}–C_{22}). These compounds are most likely from a marine source ($C_{14:0}$, $C_{16:0}$, $C_{16:1}$, and $C_{18:1}$)[18], although terrestrial plant wax bacteria, present on plant wax surfaces, also contain these components[19–21]. Marine-derived fatty acid salt and ester (C_{13}–C_{22}) and sea-salt concentrations are lower in July than in April and May (Table 1, ref. 22). This may be due to the 50% decrease in mean wind speed in July relative to the springtime.

The lower molecular weight portion of the gas chromatogram for the aliphatic hydrocarbons and fatty alcohols had few peaks other than those described above. However, several unidentified low molecular weight fatty acids (C_{13}–C_{22}) were present in addition to the saturated *n*-fatty acids. The structures of these compounds await GC–MS elucidation, although it has been determined that polyunsaturated acids are <20% in either fraction. Again, the low level of unsaturated compounds is not surprising considering their oxidative and photochemical lability relative to their saturated counterparts. The fatty acid results from Table 1 show that in many cases the salts are >10 times higher in concentration than the esters. To our knowledge, these are the first results comparing these two fatty acid fractions from the same samples.

The data shown in Table 1 and Fig. 1 clearly show a strong terrestrial plant surface wax input for the lipids analysed. Finding a strong input of terrestrial plant waxes >5,000 km from a continental source was certainly unexpected. However, mineral dust, carbon isotope and radionuclide studies yielded results explaining both the presence of terrestrial plant surface wax lipids and their decrease with time at the Enewetak station. Duce *et al.*[22] used the atmospheric Al concentration as an indicator of continental dust produced from soil and crustal weathering processes. A dramatic decrease of Al concentration occurred from mid-April to early August. Although the concentrations of three lipid class compounds do not follow this decrease exactly, the same general trend is present. The lack of an exact correlation is not surprising because the particles containing these lipids, plant epicuticular waxes, soil or loess, probably have somewhat different terrestrial sources than the dust alumino-silicates. Similarly, the ^{210}Pb activity of the particulates collected during the Enewetak experiment dropped markedly during the five month field program (K. Turekian, personal communication), again roughly in parallel with the total lipid concentrations. ^{210}Pb derives ultimately from a gaseous land source (^{222}Rn) but is formed and attaches to particles in the atmosphere. The rough similarity of these trends can be understood in terms of a fine particle, land-derived source for the organic compounds, transported offshore, initially in a continental (^{222}Rn) air mass. Carbon isotopic studies[23] also showed a continental source for the small size particles (<0.5 μm). These investigators found $\delta^{13}C$ ratios of −26 to −28‰ for four samples collected during the Enewetak experiment.

The decrease in dust concentration at Enewetak during the five month experiment was rationalized on the basis of the seasonality of dust storm activity in China and the seasonal changes in the large scale wind fields over the Pacific[22]. This observation of such large amounts of Asian dust during April and May at Enewetak was unexpected because the surface winds are within the easterly (trade) wind regime. However, a strong westerly flow north of 20° N extending from Asia into the central North Pacific is present at 700 mbar (3,000 m). During June–August, the latter part of the Enewetak experiment, conditions were not favourable for the transport of dust to Enewetak due to a decrease in these upper altitude winds[22].

The concentrations of the particulate lipid class compounds reported here are, to our knowledge, the lowest yet observed for open ocean areas far from continental influence. Unfortunately the paucity of other work and differences in methodology make precise comparisons with other regions difficult. The most comparable data are those of Hahn and coworkers (refs 7, 10 and personal communication) who found concentrations of individual particulate *n*-alkanes in the C_{21}–C_{28} region in the range of 200–550 pg m^{-3} with no odd/even carbon preference for marine aerosol samples collected off the west coast of Ireland. Similar levels were indicated for clean marine air at Cape Grim, Tasmania and in four other samples collected from ships in North Atlantic air masses (refs 7, 10 and J. Hahn, personal communication). Unfortunately, as there are no data for compounds above nC_{28}, any marked odd/even alternation, strongly indicative of land source (Fig. 1*a*), cannot be discerned.

Fatty alcohol, fatty acid and *n*-alkane values have been reported for dusts collected using large nets off the West African coast[6]: a several orders of magnitude concentration range was observed. The lower values found, for samples collected in the south-east trades, were in the same concentration range as those found in Enewetak (Table 1), and the individual compound distributions of the lipids were very similar. However, the net dust data are only for large particles (>2 μm diameter) while the Enewetak data are for all particles >0.03 μm diameter. Marty *et al.*[8] reported much higher aerosol *n*-alkane values of 30–50 ng m^{-3} and fatty acid values of 80–230 ng m^{-3} from samples collected off north-west Africa on glass fibre filters. Their *n*-alkane data show evidence of fossil fuel input and no evidence for a terrestrial plant source fraction. Barger and Garrett[1,3] analysed aerosol samples from the eastern equatorial and North central Pacific for fatty acids in the C_{12}–C_{20} range. As no data are reported for compounds >C_{20} the fatty acid land plant component of these samples cannot be assessed.

Our preliminary data therefore show the presence of high molecular weight lipid class compounds in marine aerosols at a remote open ocean site.

The individual compound distribution data strongly indicate a continental source, probably epicuticular vascular plant waxes, for a significant portion of this material. The temporal variation, a decreasing trend in alumino-silicates, ^{210}Pb and total lipids with time, strongly suggest that similar transport mechanisms and source functions may be involved. Further verification of these interpretations requires size-fractionated sampling (by cascade impactor), continued maintenance of ultra-clean sampling conditions, associated meteorological controls and other chemical analyses to infer sources and transport processes from lipid concentration data.

We thank the Oceanography Section of the NSF for support (OCE 77-12914) as part of the SEAREX Program and R. A. Duce, C. Lee and J. Volkman for critical comments. We also thank J. Alford, E. Atlas, R. Cayer, C. Snipes, the staff of the University of Hawaii's Mid-Pacific Marine Laboratory and the Department of Energy for support during the field work in Enewetak.

Received 7 January; accepted 9 March 1981.

1. Barger, W. R. & Garrett, W. D. *J. geophys. Res.* **75,** 4561–4566 (1970).
2. Quinn, J. G. & Wade, T. L. in *Baseline Studies of Pollutants in the Marine Environment* (ed. Goldberg, E. D.) 633–663 (National Science Foundation, Washington DC, 1972).
3. Barger, W. R. & Garrett, W. D. *J. geophys. Res.* **81,** 3151–3157 (1976).
4. Ketseridis, G., Hahn, J., Jaenicke, R. & Junge, C. *Atmos. Envir.* **10,** 603–610 (1976).
5. Cautreels, W., Van Cauwenberghe, K. & Guzman, L. A. *Sci. Total Envir.* **8,** 79–88 (1977).
6. Simoneit, B. R. T. *Mar. Chem.* **5,** 443–464 (1977).
7. Eichmann, R. *et al. Atmos. Envir.* **13,** 587–599 (1979).
8. Marty, J. C., Saliot, A., Buat-Menard, P., Chesselet, R. & Hunter, K. A. *J. geophys. Res.* **84,** 5707–5716 (1979).
9. Van Vaeck, L., Broddin, G. & Van Cauwenberghe, K. *Envir. Sci. Technol.* **13,** 1494–1502 (1979).
10. Eichmann, R., *et al. Atmos. Envir.* **14:** 695–703 (1980).
11. Simoneit, B. R. T. in *Advances in Organic Geochemistry* (eds Douglas, A. G. & Maxwell, J. R.) 343–352 (Pergamon, Oxford, 1980).
12. Peltzer, E. T., Gagosian, R. B., Zafiriou, O. C. & Alford, J. (in preparation).
13. Lockey, K. H. *Comp. Biochem. Physiol.* **65B,** 457–462 (1980).
14. Grob, K. & Grob, G. *J. High Res. Chrom. Chrom. Commun.* **3,** 197–198 (1980).
15. Eglinton, G. & Hamilton, R. J. in *Chemical Plant Taxonomy* (ed. T. Swain) 187–217 (Academic, New York, 1963).
16. Tulloch, A. P. in *Chemistry and Biochemistry of Natural Waxes* (ed. Kolattukudy, P. E.) 236–252 (Elsevier, New York, 1976).
17. Farrington, J. W. in *Advances in Chemistry*, Ser. No. 185. (eds Petrakis, L. & Weiss, F. T.) 1–22 (American Chemical Society, Washington DC, 1980).
18. Schultz, D. M. & Quinn, J. G. *J. Fish. Res. Bd Can.* **29,** 1482–1486 (1972).
19. Simoneit, B. R. T. in *Proc. Carbonaceous Particles in the Atmosphere* (ed. Novakov, T.) 233–244 (National Science Foundation, Lawrence Berkeley Laboratory, 1979).
20. Volkman, J. K., Johns, R. B., Gillan, F. T., Perry G. J. & Bavor, H. J. Jr *Geochim. cosmochim. Acta* **44,** 1133–1143 (1980).
21. Smith, W. H. & Staskawicz, B. J. *Envir. Manag.* **1,** 317–330 (1977).
22. Duce, R. A., Unni, C. K., Ray, B. J., Prospero, J. M. & Merrill, J. T. *Science* **209:** 1522–1524 (1980).
23. Chesselet, R., Fontugne, M., Buat-Menard, P., Ezat, U. & Lambert, C. E. *Geophys. Res. Lett.* (in the press).

Floristic changes indicate a cooling climate in the Eocene of southern England

M. E. Collinson*, K. Fowler† & M. C. Boulter‡

* Palaeontology Department, British Museum (Natural History), Cromwell Road, London SW7 5BD, UK
† Department of Biological Sciences, Portsmouth Polytechnic, King Henry I Street, Portsmouth PO1 2DY, UK
‡ Department of Biological Sciences, North East London Polytechnic, Romford Road, London E15 4LZ, UK

A major climatic cooling at the close of the Eocene (the 'terminal Eocene event') is suggested by studies of leaf floras in North America[1,2]; and by oxygen isotope studies, of mollusc shells from the North Sea[3] and of planktonic and benthic foraminifera from the North and South Pacific[4]. Climatic fluctuations in the French Palaeogene[5–8] inferred from palynological studies, are difficult to interpret in terms of this single event. The Pacific faunal studies[4] and preliminary palynological evidence from the Isle of Wight, southern England[9] indicates an onset of cooling earlier in the Eocene. We have therefore examined fossil floras (fruits, seeds, pollen and spores) from the Palaeogene deposits of the London and Hampshire Basins of southern England which form a continuous sequence from the early Eocene. We find no evidence for a sudden climatic change at the end of the Eocene (taken to be near the base of the Bembridge Marls in the Hampshire Basin[10]). Our evidence, in the form of two major periods of floristic change, suggests a more gradual cooling, commencing in the latest early Eocene.

The evidence of floristic change during the Eocene has resulted from recent work on fossil pollen and spores from coastal sections (Hampshire Basin[9]) and Institute of Geological Sciences boreholes (Hampshire Basin[11]) together with fossil angiosperm fruits and seeds (throughout the area[12]). Palynological evidence is available as percentage frequency counts from some 200 samples representing two Eocene sections in the Hampshire Basin. The sections were at Whitecliff Bay, Isle of Wight (SZ 643 864) which includes the type section of the Bracklesham Beds[13], and the I.G.S. Ramnor Inclosure borehole near Brockenhurst (SU 3114 0475). Of more than 100 form-genera present within the assemblages, many occur throughout the Hampshire Basin Tertiary and are known to range throughout the Eocene of North-west Europe. Of particular significance are those form-genera with a restricted stratigraphical range within the sections studied (Fig. 1). In the Whitecliff Bay section, 13 form-genera first appear within the Bracklesham Group and one other in the Lower Headon Beds; 11 extinctions occur in the lower part of the Bracklesham Group and another six at the end of the Lower Headon Beds. In the Ramnor borehole, seven form-genera first appear within the Barton and Headon Beds and eight become extinct within the Bracklesham Group.

Among the fossil fruits and seeds, 72 angiosperm taxa have a stratigraphical range within the Palaeogene of southern England (data from the same sources as those on Fig. 2). Of these taxa, 45% first appear within the London Clay and 13% within the Bracklesham Group. Major extinctions occur within the Bracklesham Group (14%) and the Lower Headon Beds (39%).

The sedimentary sequence which we are studying was deposited in a marginal area affected by marine transgressions and regressions. It is recognized that facies variation in such a depositional environment could account for the apparent macrofloristic changes. Therefore, detailed analyses are presented for 35 taxa (including 22 already considered above) all of which have their nearest living relatives occupying a fresh or brackish, aquatic or marginal aquatic habitat (Fig. 2). Lithofacies representing deposition in these habitats occur extensively throughout the study area, whereas those containing land-derived macrofossils have a more sporadic occurrence. Macrofloristic data in Fig. 2 are considered to represent a close approximation to the true distribution of the plants which bore these fossil fruits and seeds. This is further supported by evidence that many of the fossil occurrences cross facies boundaries within the sequence. The macro- and microfloral changes which

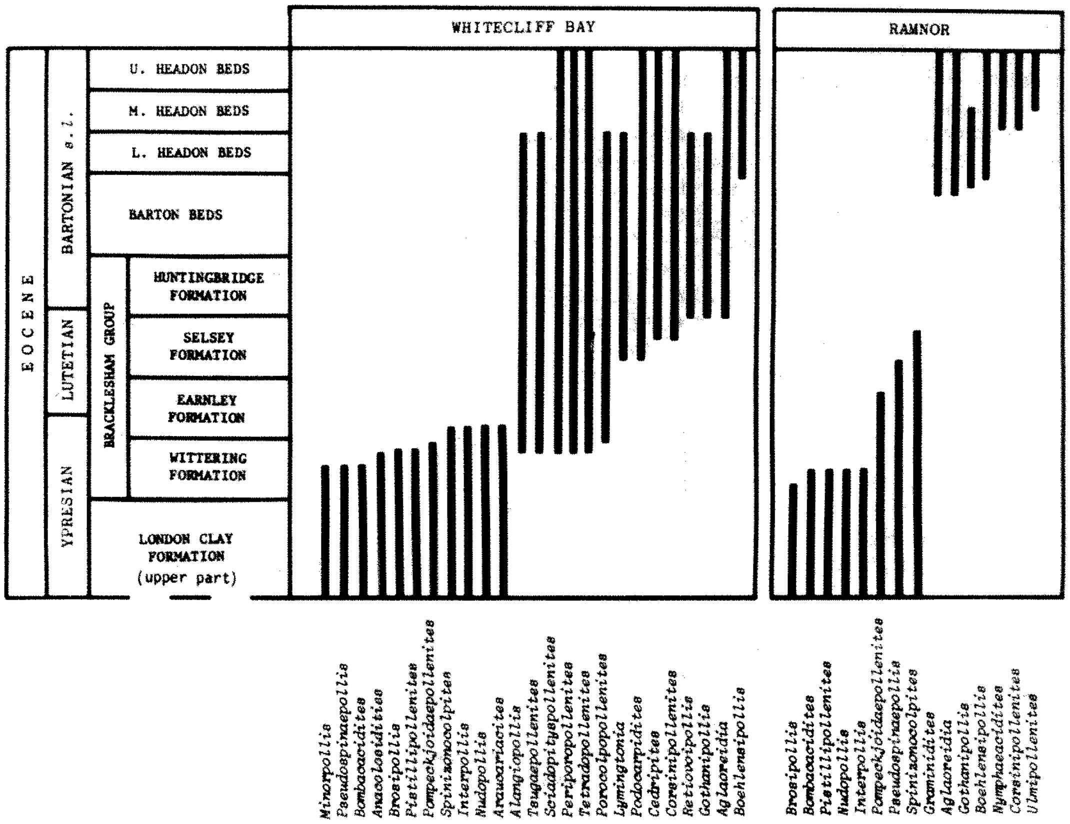


Fig. 1 Stratigraphical ranges of selected pollen and spore form-genera in the Eocene of the Whitecliff Bay coastal section and Ramnor Inclosure borehole. Stratigraphy follows that of Curry *et al.*[10] and European stage names extrapolated from refs 10, 19.

0028–0836/81/220315—03$01.00

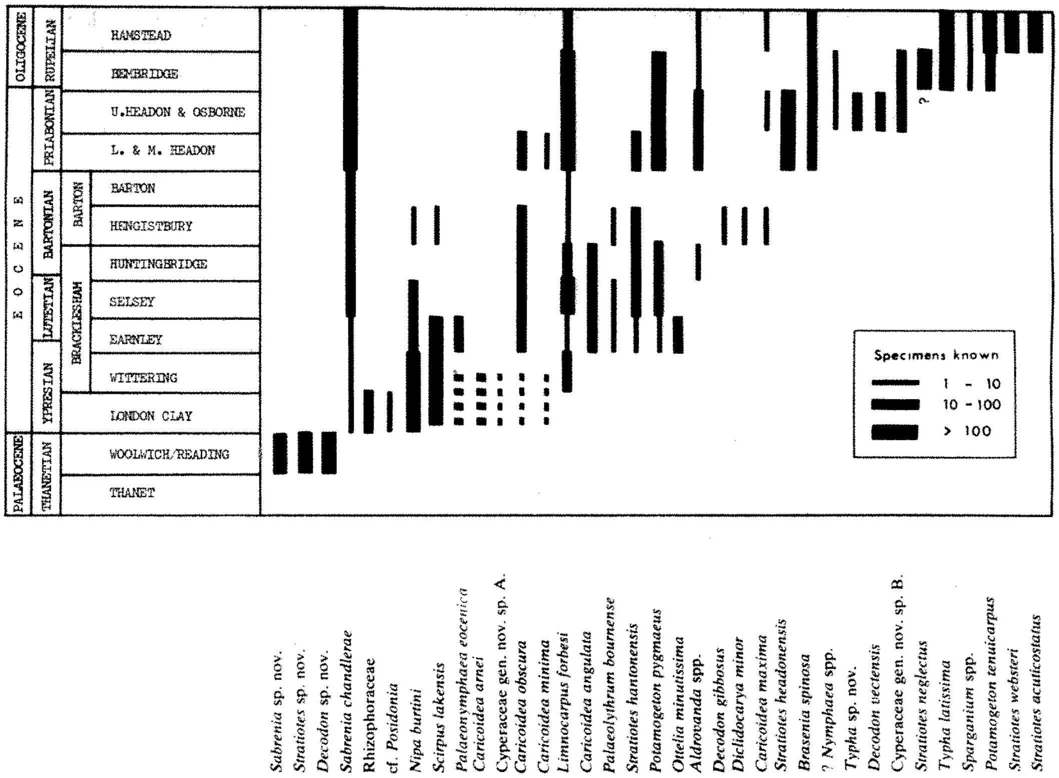


Fig. 2 Stratigraphical distribution of selected fossil angiosperm fruits and seeds in the Palaeogene of southern England. Stratigraphy follows Curry *et al.*[10] and the lithostratigraphical sequence of Hooker *et al.*[20] Stage names have been positioned by extrapolation from refs 10, 19. Macrofloral data sources are from the National collections and refs 12, 21, 22. Plants from the Cuisian of Bracklesham Bay[23] are here considered to be Wittering in age, those from the Lutetian[23] to be Earnley and those from the Auversian and Selsey[23] to be Selsey. Plants from the Alum Bay plant bed[24] (Prestwich Bed 24) are here considered to be Earnley; those from the Alum Bay plant bed[24] (Prestwich Bed 17) and Fisher[13] Bed V in Whitecliff Bay to be Wittering in age. Various localities of Dorset Pipe Clays and Bagshot Beds[12,25] have been placed within the London Clay and Wittering. Short dashed lines indicate uncertainty within this range. The positions of other plant bearing strata may be extrapolated from ref. 10.

we recognize do not coincide with major facies change. Despite facies differences between Whitecliff Bay and Alum Bay, Isle of Wight, which reflect differing depositional environments, essentially the same microfloristic changes were observed at both localities.

Three main assemblages may be recognized within the fruit and seed floras. The first occurs in the London Clay and lower Bracklesham Group. It is characterized by mangrove taxa related to modern *Nipa* and the Rhizophoraceae, reed swamp elements such as *Scirpus lakensis* and other Cyperaceae, and open water plants represented by *Palaeonymphaea* and *Sabrenia*. Pollen with a comparable stratigraphical range include *Spinizonocolpites*, *Bombacacidites*, *Compositoipollenites*, *Anacolosidites* (showing affinities with modern tropical rain forest types), *Pistillipollenites*, *Brosipollis*, *Interpollis*, *Nudopollis*, *Pseudospinaepollis*, *Pompeckjoidaepollenites* and *Minorpollis* (considered to belong to extinct groups).

A significant microfloristic change occurs in the lower Bracklesham Group at Whitecliff Bay marked by the disappearance of taxa typical of the earlier Eocene. During Bracklesham time the palynoflora shows progressive change with many introductions and extinctions. The dominant elements of the first 'aquatic' macrofloral assemblage are gradually replaced during Bracklesham times by a second assemblage with seeds of other Cyperaceae and *Limnocarpus*, *Stratiotes* and *Potamogeton*. Previously abundant tropical and subtropical macrofloral elements, such as Menispermaceae, Icacinaceae and Anonaceae, are much reduced in number by late Bracklesham time. Both these and the dominant lower Bracklesham aquatic plants become absent from the sequence by the close of Lower Headon Beds deposition.

In the upper Bracklesham of Whitecliff Bay a second sequence of microfloristic changes begins. It is marked by the introduction of new form-genera characteristic of the overlying Barton and Headon Beds, together with an increased abundance of spores, bisaccate and inaperturate pollen (representing ferns and conifers). This fern–conifer association is a dominant feature of palynological assemblages in the overlying Palaeogene sediments of the Hampshire Basin. Further changes occur in the microflora in the Lower and Upper Headon Beds, and a third macrofloral assemblage originates at this time with additions in the Bembridge and Hamstead Beds. *Typha* emerges as the dominant reed swamp element and open water aquatic plants are characteristic.

Of the restricted pollen form-genera, the morphologically distinctive *Spinizonocolpites* is clearly related to the modern tropical genus *Nipa*. This pollen becomes absent from the sequence at approximately the same time as the major fall and hiatus in the record of fossil *Nipa* fruits. These latter do not disappear from the sequence until slightly later. This stratigraphical distribution of *Nipa* fruits and *Spinizonocolpites* pollen is similar to that found elsewhere in Europe. The pollen reaches maximum development in the early Middle Eocene[14,15] and the fruits attain their maximum size during the same period in Britain, Belgium and France.

Each line of evidence indicates an important floristic change during deposition of the Bracklesham Group. Extinctions in the microflora in the lower part of the Bracklesham Group have not been recorded in the fruit and seed floras. Otherwise, first appearances and extinctions during Bracklesham time are common to the micro- and macrofloras. Optimal development of both fossil pollen and fruit related to the living tropical *Nipa* occurs during middle Bracklesham time. Floristic change also occurs within the late Barton and Headon Beds. Major extinctions in the macroflora and changes in the microflora of the Lower Headon Beds may be related to unstable edaphic conditions reflected by the onset of fluvio-marine sedimentation[16].

These results indicate two main periods of floristic change during the Eocene in Britain. The first originated in the Wittering and continued during the Lutetian. The second originated in the uppermost Selsey and is expressed mainly within the Barton and Headon Beds. This second change culminates in the development of extensive reed swamp floodplain vegetation[12,17], the extinction of many tropical and subtropical floral elements and the dominance of pteridophyte spores and conifer pollen in the microfloral assemblages. Both of these floristic changes may indicate a progressive lowering of land surface temperature during the Middle Eocene in southern

England, the second possibly more rapid than the first. The cooling which we suggest these changes may imply took place over a period of 15 Myr. This is of interest in relation to the palaeotemperature curves for the North Sea[3,18], which indicate a sudden and rapid cooling at the close of the Eocene. The floristic changes shown in our results from the Palaeogene of southern England suggest that climatic events were more complex than is indicated by the recognition of a single 'terminal Eocene event'.

M.E.C. acknowledges receipt of a NERC research fellowship during the tenure of which part of this work was completed. Results from the Ramnor borehole are published with permission of the director of Institute of Geological Sciences.

Received 7 January; accepted 10 March 1980.

1. Wolfe, J. A. *Am. Scient.* **66,** 694–703 (1978).
2. Wolfe, J. A. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **30,** 313–323 (1980).
3. Buchardt, B. *Nature* **275,** 121–123 (1978).
4. Savin, S. M. *A. Rev. Earth planet. Sci.* **5,** 319–355 (1977).
5. Gruas-Cavagnetto, C. *Mém. Soc. géol. Fr.* **131,** 1–64 (1978).
6. Cavelier, C. *Sci. géol. (Strasbourg) Mém.* **54,** 1–280 (1979).
7. Chateauneuf, J-J. *Mém. Bur. Rech. géol. Minièr.* **116,** 1–360 (1980).
8. Ollivier-Pierre, M-F. *Mém. Soc. géol. minér. Bretagne* **25,** 1–239 (1980).
9. Fowler, K. in *5th int. Palyn. Conf.,* Cambridge (1980).
10. Curry, D. *et al. Geol. Soc. Lond. spec. Rep.* **12,** 1–72 (1978).
11. I. G. S. boreholes 1977. *Rep. Inst. geol. Sci.* No. 78/21 (1978).
12. Collinson, M. E. thesis, Univ. London (1978).
13. Fisher, O. *Q. Jl geol. Soc. Lond.* **18,** 65–94 (1862).
14. Chateauneuf, J. J. *Bull. B. R. G. M. Sect. IV, Géol. Générale. Suppl. Serv. Géol. Nat.* RS 812 (1979).
15. Haseldonckx, P. *Geol. Mijnb.* **51,** 645–650 (1972).
16. Fowler, K. *Pollen Spores* **13,** 135–147 (1971).
17. Fowler, K., Edwards, N. & Brett, D. *Palaeontology* **16,** 205–217 (1973).
18. Buchardt, B. *Dansk. Met. Inst. Climatol. Pap.* **4,** 243–248 (1978).
19. Hardenbol, J. & Berggren, W. A. *A.A.P.G. Stud. Geol.* **6,** 213–234 (1978).
20. Hooker, J. J., Insole, A. N., Moody, R. T. J., Walker, C. A. & Ward, D. J. *Tert. Res.* **3,** 1–45 (1980).
21. Chandler, M. E. J. *The Lower Tertiary Floras of Southern England* Vol. 4 (British Museum Natural History, London, 1964).
22. Collinson, M. E. *Ann. Bot.* **46,** 603–632 (1980).
23. Chandler, M. E. J. *Bull. Br. Mus. nat. Hist. geol.* **5,** 13–41 (1961).
24. Crane, P. R. *Tert. Res.* **1,** 95–99 (1977).
25. Chandler, M. E. J. *The Lower Tertiary Floras of Southern England* Vol. 2 (British Museum Natural History, London, 1962).

A new assessment of *Rhyniella*, the earliest known insect, from the Devonian of Rhynie, Scotland

Paul Whalley & E. A. Jarzembowski

British Museum (Natural History), Department of Entomology,
Cromwell Rd, London SW7 5BD, UK

The search for insects (Class Hexapoda) in rocks of Devonian and Lower Carboniferous age (325–415 Myr BP) has continued for many years because so little is known of the early history of these diverse and abundant animals. In 1926, four head capsules of an insect, named *Rhyniella praecursor*, were discovered in Lower Devonian chert from Rhynie, Aberdeenshire[1] by examining thin fragments of this chert, which are rather translucent, like amber. Subsequently, about 10 additional specimens were found showing the legs and thorax of this insect[2]. It has become generally accepted that *R. praecursor* belongs to the living Order Collembola or springtails, a widespread, flightless group of insects. However, several characters of Collembola were not preserved in the fossils and doubts remained. Furthermore, the apparent similarity of *Rhyniella* to some specialized modern Collembola[3] has suggested that *Rhyniella* was a modern contaminant[4]. Our research shows that *Rhyniella* is not a contaminant; we have also found the first intact abdomen of *Rhyniella* which has a special locomotory appendage similar to that of modern Collembola.

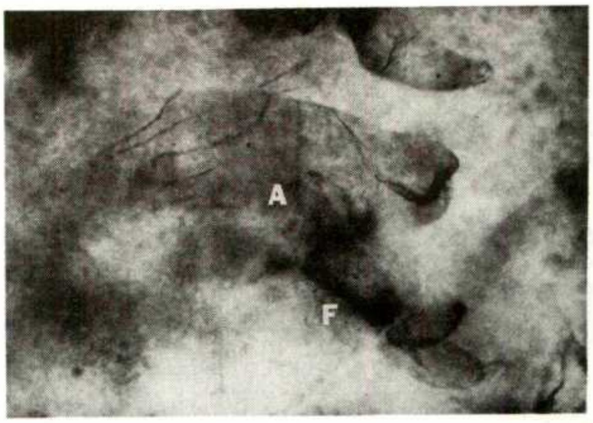


Fig. 1 Photograph of *R. praecursor,* fossilized in chert from Rhynie. A, abdomen; F, furcula.

We initially decided to search the Rhynie Chert because it is the only formation before the Upper Carboniferous to have yielded insect remains. The samples we examined were excavated during the International Botanical Congress in 1964 by J. Pettitt and C. Shute of the Department of Palaeontology, British Museum (Natural History). However, no insects were found although several arachnids were discovered. We then re-examined the original material collected over 50 yr ago; in a few cases the chert was thick and thus was ground and repolished to see the contents more clearly. As a result, a new specimen was revealed which consists of a thorax with legs and the first-discovered complete abdomen. Extending from the ventral surface of the abdomen is a structure corresponding to the furcula ('spring') of Recent Collembola (Figs 1, 2). Only the hind end of the *Rhyniella* abdomen can be seen in Fig. 1 because the body is oblique to the plane of view. The presence of a furcula confirms that *Rhyniella* was a springtail; the total length of the insect was ~1.5 mm. Figure 2 shows a Recent springtail which has a slender apex (dentes) to the furcula.

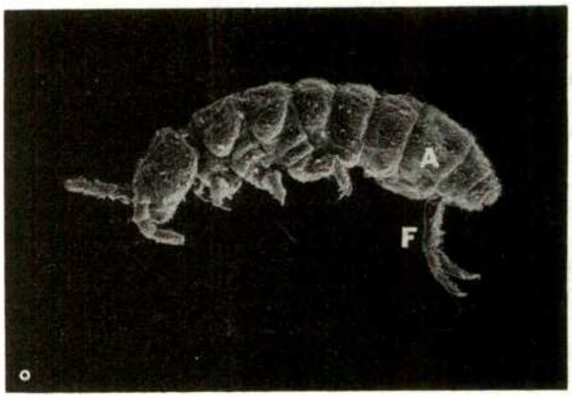


Fig. 2 A Recent springtail. A, abdomen; F, furcula.

Examination of several specimens of *Rhyniella* showed that they were completely fossilized in chert and did not look like modern contaminants. They were examined petrologically in polarized light by A. R. Woolley and the enclosing chert was found to be homogeneous and certainly represented a single phase of mineral growth. One *Rhyniella* specimen was associated with a plant fragment in homogeneous chert and J. B. Richardson confirmed that the petrification of this was identical to other Rhynie plants of undoubtedly Devonian age. We conclude that *Rhyniella praecursor* is of Lower Devonian age (possibly Siegenian).

Thus Collembola were established and already developing their specialized mode of locomotion some 380 Myr ago. The relationship of *Rhyniella* with Recent Collembola Neanuridae is doubtful and a fuller account will be published elsewhere.

We thank P. N. Lawrence for advice on Collembola.

Received 24 November 1980; accepted 30 March 1981.

1. Hirst, S. & Maulik, S. *Geol. Mag.* **63,** 69–71 (1926).
2. Scourfield, D. J. *Nature* **145,** 799–804 (1940).
3. Massoud, Z. *Revue Ecol. Biol. Sol.* **4**(3), 497–505 (1967).
4. Crowson, R. A. *Classification and Biology,* 1–350 (Heinemann, London, 1970).

0028-0836/81/220317—01$01.00

The role of GABAergic inhibition in the cortical effects of monocular deprivation

Adam M. Sillito & John A. Kemp

Department of Physiology, Medical School,
Birmingham B15 2TJ, UK

Colin Blakemore

University Laboratory of Physiology, Parks Road,
Oxford OX1 2JD, UK

In cats and monkeys, monocular deprivation in early life results in an extensive reduction of the capability of the deprived eye to activate visual cortical cells[1], apparently because of changes taking place in the visual cortex itself rather than at an earlier stage in the visual pathway[2]. The nature of the change taking place in the cortex is the subject of some controversy. One line of evidence favours a redistribution of the excitatory terminals relating to the two eye inputs[3-5] whilst another suggests that the deprived eye input is in some way suppressed[6], possibly by γ-aminobutyric acid (GABA)-mediated intracortical inhibition[7]. We have now sought to clarify the role of this type of inhibitory influence in monocularly deprived cats, by ascertaining whether there is an enhancement of the deprived eye input to visual cortical cells during the iontophoretic application of the GABA antagonist, *N*-methyl-bicuculline (NMB). Our experiments show that although intracortical inhibition may submerge the deprived eye input to some cells, it is unlikely to precipitate the cortical effects of monocular deprivation. The redistribution of excitatory terminals would seem to be a more significant factor.

Six cats were used for the experiments—three early monocularly deprived animals, with lid suture at the time of natural eye opening and three late monocularly deprived animals, with lid suture carried out 2 weeks after eye opening. All recordings were made when the animals were in the age range 12–24 months. After initial surgical preparation the animals were paralysed with Flaxedil and maintained on an anaesthetic mixture of 75% N_2O–25% O_2 and 0.1–0.4% Halothane. Full details of experimental procedures are given elsewhere[8]. Recordings were made extracellularly from the central barrel of five-barrel micropipettes, the other barrels containing respectively NMB (two barrels), DL-homocysteate and GABA. Cell properties were documented on line by construction of peristimulus time histograms (PSTHs) of the responses to a slit of light moving over the receptive field in the normal eye and the corresponding location in the deprived eye. Ocular dominance was assessed before, during and after NMB application. Retinal landmarks were periodically back-projected onto the stimulus projection screen, to check the relative position of the two eyes. Great care was taken in varying stimulus location and orientation in the deprived eye during NMB application to ensure that errors in estimation of corresponding field location and eye rotation did not cause weak response fields to be missed.

Tests were completed on 51 cells, all initially exclusively dominated by the normal eye. Of these cells, 29% (15) developed a response, though sometimes fairly weak, to stimulation of the deprived eye during NMB application. The results for the early and late monocularly deprived animals were essentially the same, with respectively 28% (6/21) and 30% (9/30) of the cells studied in each group developing deprived eye inputs. The pooled results for early and late monocularly deprived animals are summarized in Fig. 1. These should be compared with similar data obtained for monocularly dominated cells in normal cats[9], where ~50% of the cells showed a change in ocular dominance during NMB application. This is a significantly larger percentage than that seen in the present sample of cells and suggests that intracortical inhibition exerts a more powerful influence on ocular dominance in the cortex of normal as opposed to monocularly deprived cats.

An example of the action of NMB on the ocular dominance of a complex cell is illustrated in Fig. 2. This shows the effect of two successive periods of NMB application on the response to movement of an optimally oriented slit of light over the normal eye's field and the corresponding location in the deprived eye. During each period of NMB application there was a loss of directional selectivity and an increase in magnitude of the normal eye response, together with the appearance of a smaller but clear response in the deprived eye. These changes were followed each time by recovery of the original properties after termination of the application. The results for all the cells developing a deprived eye input were reversible and

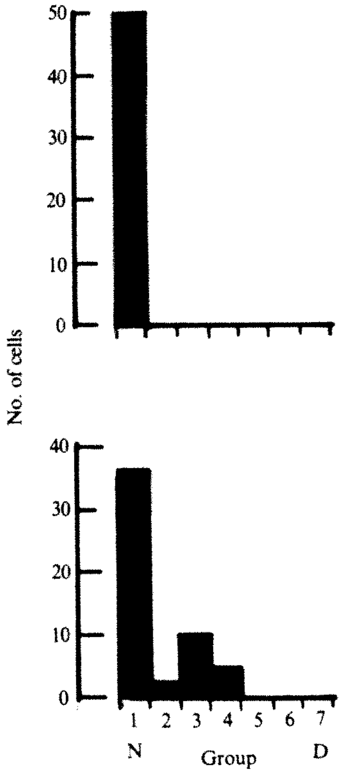


Fig. 1 Histograms showing the effect of iontophoretic application of NMB on the ocular dominance of cells in the visual cortex of six monocularly deprived cats. The ocular dominance of the cells has been classified on the seven-point scale described by Hubel and Wiesel[15]. Group '1' cells are exclusively driven by the contralateral eye, group 7 by the ipsilateral eye, while those in group 4 are equally driven by both eyes. Cells in groups 5 and 6, and 3 and 2, are slightly and strongly dominated by ipsilateral and contralateral eyes, respectively. Recordings were made in area 17 both ipsilateral and contralateral to the deprived eye. All the cells studied were initially exclusively dominated by the normal eye and hence fell into group 1 or 7. In the histograms, the results from ipsilateral area 17 have been pooled with those from contralateral area 17 so that group 1 represents the normal eye (N) whilst group 7 represents the deprived eye (D). A shift in ocular dominance from left to right represents a change in favour of the deprived eye. The upper histogram shows the control ocular dominance distribution and the lower histogram the distribution during NMB application. The largest shift in ocular dominance was to group 4 with equal driving by normal and deprived eye.

0028-0836/81/220318—03$01.00

reproducible. We also observed that the response characteristics obtained from stimulation of the normal and deprived eyes were often quite distinct. However, we are not in a position to draw any conclusions regarding possible differences in the overall response specificity exhibited by the latent deprived eye receptive fields and the normal eye fields[6]. The problem is that many of the receptive field properties of visual cortical cells, including orientation selectivity, are partially or entirely dependent on the action of GABA-mediated intracortical inhibition[8,10–12]. Consequently, apart from any action on ocular dominance, iontophoretic application of NMB also influences those processes generating or enhancing receptive field specificity. Thus, we cannot say whether a lack of specificity in the deprived eye input reflects the action of NMB on GABA-mediated inhibitory mechanisms that generate the specificity, or the absence of such mechanisms operating with respect to the deprived eye input. The report[7] that intravenous injection of bicuculline revealed deprived eye fields with normal properties in monocularly deprived cats is puzzling and would seem to indicate that the intravenous bicuculline was having little effect on intracortical GABAergic synapses.

One of the side effects of NMB application is a marked elevation in resting discharge level. This may be partly due to release of the cell from tonic GABAergic inhibitory effects and partly to a direct excitatory action of the NMB. The lower record in Fig. 2 shows an attempt to control for the possible effects of the elevated resting discharge by generating an increased discharge level by the iontophoretic application of an excitatory amino acid, DL-homocysteate. In these circumstances, at a similar resting discharge level to that seen with NMB, there is no sign of a visual response from the deprived eye. This suggests that in these experiments it is the block of GABAergic inhibition and not the increase in resting discharge as such that reveals the deprived eye input. Indeed, it seems that, for many visual cortical cells, an increase in resting discharge in the absence of a block of GABAergic inhibition decreases responsiveness to visual stimuli and increases the effectiveness of visually evoked inhibitory influences, possibly via recurrent collateral feedback[8]. It is interesting that Blakemore, Hawken and Mark (in preparation) have shown that iontophoretic appliation of DL-homocysteate at currents sufficient to depolarize cells without causing considerable spontaneous discharge will reveal an input from the deprived eye to many neurones in briefly monocularly deprived kittens.

In experiments of this type negative results must be interpreted with greater caution than positive ones. The fact that 70% of the cells failed to show an input from the deprived eye during NMB application suggests that intracortical GABA-mediated inhibition is making no significant contribution to their domination by the normal eye. However, it could reflect a lack of pharmacological effectiveness of the NMB or an inadequate concentration of the drug at the actual inhibitory synapses involved. Control observations carried out on all cells included in the present sample showed that these possibilities were not significantly influencing the results. We determined the ejecting current necessary for iontophoretically applied GABA to suppress completely the cell's response to motion of an optimal stimulus over the normal eye receptive field. Data taken during NMB application were only considered acceptable when the previously determined suppressive dose of GABA failed to suppress the cells driven activity. This is illustrated in Fig. 3 for a complex cell that retained its ocular dominance during NMB application. The PSTHs show responses respectively to normal eye stimulation, deprived eye stimulation and then normal eye stimulation again but in the presence of iontophoretic application of GABA. During NMB application, although there is no response from the deprived eye, the application of GABA fails to suppress the normal eye response, indicating that the NMB is pharmacologically effective. With this type of pharmacological check and other functional criteria, we assessed the potency of NMB at the synaptic level. For all cells, NMB application resulted in a reduction or loss of inhibition-dependent receptive field properties[8,10–12] such as direction and orientation selectivity, together with an increased magnitude of response suggesting a loss of tonic inhibitory influences. These observations indicate that neither variations in the pharmacological potency of NMB nor its functional effectiveness at inhibitory synapses were significant factors contributing to the failure to reveal deprived eye receptive field properties in 70% of the cells studied. However, inhibitory influences other than those mediated by GABA could be suppressing the deprived eye input to some cells. Although strong evidence supports the role of

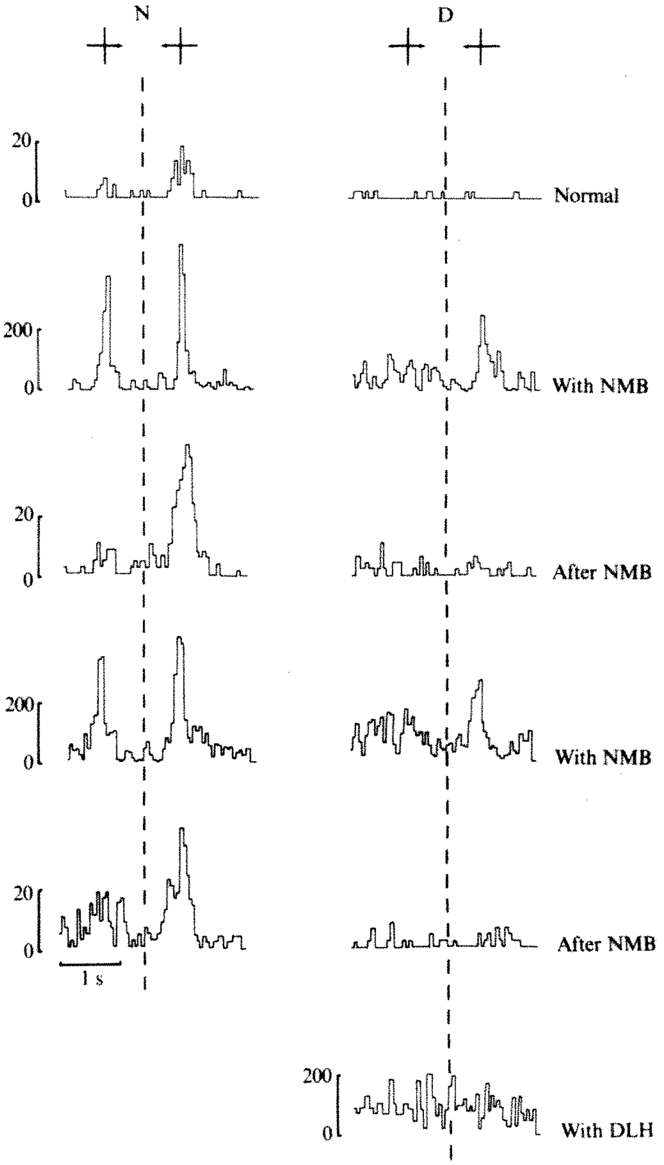


Fig. 2 Responses of a complex cell showing a shift in ocular dominance during iontophoretic application of NMB. Each PSTH shows the responses to 25 sweeps of an optimally oriented slit of light moving forwards (to left of dotted line) and backwards (to right of dotted line) over the receptive field in the normal eye (N) and the corresponding location in the deprived eye (D). Bin size 50 ms. Vertical calibration indicates the number of spikes per bin accumulated in 25 trials. PSTHs document the effect of and recovery from two successive periods of the iontophoretic application of NMB (ejecting current 60 nA). The bottom right PSTH shows that increasing the resting discharge by iontophoretically applying an excitatory amino acid, DL-homocysteate (DLH, ejecting current 8 nA), does not restore a response in the deprived eye.

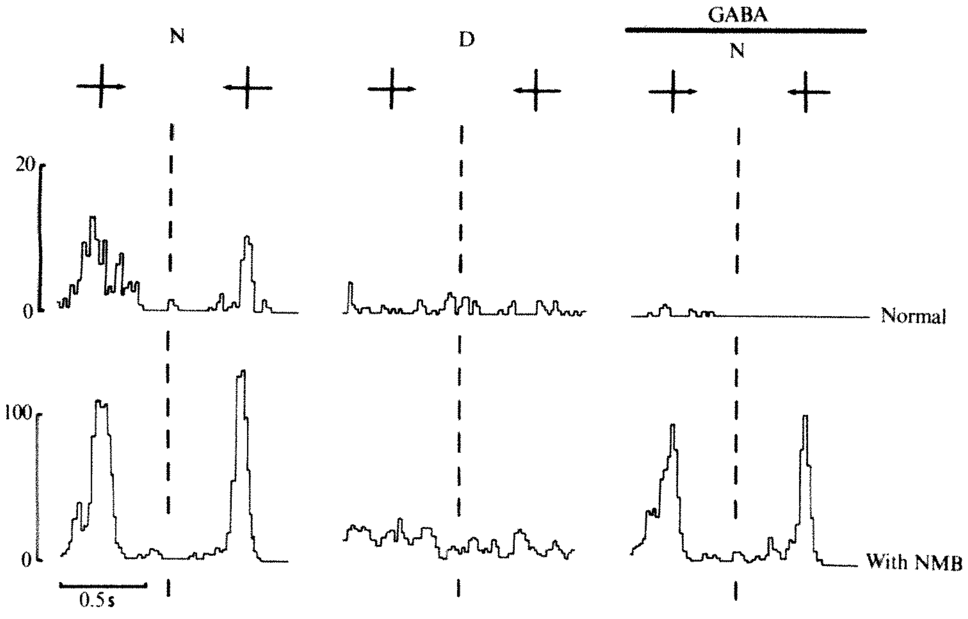


Fig. 3 Responses of complex cell showing pharmacological block of GABA action during NMB application but no deprived eye input. PSTHs show response to an optimally oriented slit moving over the normal eye receptive field, corresponding location in the deprived eye and normal eye receptive field again but during iontophoretic application of GABA (15 nA, marked by bar over records). Details of PSTHs as for Fig. 2. Upper records show response before and lower records response during NMB application (ejecting current 60 nA). Note that GABA blocks normal eye-driven response before, but not during, NMB application.

GABA as a major inhibitory transmitter in the visual cortex[10,13,14], other inhibitory transmitters could be involved in this location. It is therefore important to emphasize that the present data only refer to GABA-mediated inhibition.

We conclude that in monocularly deprived animals 30% of the cells exclusively dominated by the normal eye actually receive a deprived eye input that is normally submerged, at least in part, as a result of the action of local GABAergic inhibitory influences acting on the cells. It is not clear whether plastic changes in the GABAergic inhibitory processes *per se* are an important or even significant factor contributing to the actual shift in ocular dominance seen during monocular deprivation. Note that none of the cells studied was ever dominated by the deprived eye during NMB application. If intracortical inhibitory mechanisms in normal animals enhance the contrast between adjacent ocular dominance columns[9], these same processes initiated from the greatly extended normal eye ocular dominance columns could suppress the output from the small islands of deprived eye input known to exist in layer IV[3,6]. Thus the precipitating factor would be the redistribution of excitatory terminals from the two eyes and not the specific action of intracortical inhibitory processes. Some of our results may also reflect a very simple nonspecific increase in responsiveness due to the removal of the tonic GABAergic inhibitory influences, whereby a small deprived eye input becomes apparent because the cell is more responsive to its input. These nonspecific effects would not, however, apply to the few cells that became equally driven by both eyes. We postulate a passive, rather than an active, role for intracortical GABAergic inhibition in the effects of monocular deprivation.

Support from the MRC and Wellcome Trust is gratefully acknowledged.

Received 3 November 1980; accepted 9 March 1981.

1. Wiesel, T. N. & Hubel, D. H. *J. Neurophysiol.* **26,** 1003–1017 (1963).
2. Wiesel, T. N. & Hubel, D. H. *J. Neurophysiol.* **26,** 994–1002 (1963).
3. Blakemore, C., Garey, L. J. & Vital-Durand, F. J. *J. Physiol., Lond.* **283,** 223–262 (1978).
4. Hubel, D. H., Wiesel, T. N. & LeVay, S. *Phil. Trans. R. Soc.* **B278,** 131–163 (1977).
5. Shatz, C. J. & Stryker, M. P. *J. Physiol., Lond.* **281,** 267–283 (1978).
6. Kratz, K. E., Spear, P. D. & Smith, D. C. *J. Neurophysiol.* **39,** 501–511 (1976).
7. Duffy, F. H., Snodgrass, S. R., Burchfiel, J. L. & Conway, J. L. *Nature* **260,** 256–257 (1976).
8. Sillito, A. M. *J. Physiol., Lond.* **289,** 33–53 (1979).
9. Sillito, A. M., Kemp, J. A. & Patel, H. *Expl Brain Res.* **41,** 1–10 (1980).
10. Sillito, A. M. *J. Physiol., Lond.* **250,** 287–304 (1975).
11. Sillito, A. M. *J. Physiol., Lond.* **271,** 699–720 (1977).
12. Tsumoto, T., Eckhart, W. & Creutzfeldt, O. D. *Expl Brain Res.* **34,** 351–363 (1979).
13. Iversen, L. L., Mitchell, J. F. & Srinivasaan, V. *J. Physiol., Lond.* **212,** 519–534 (1971).
14. Ribak, C. E. *J. Neurocytol.* **7,** 461–478 (1978).
15. Hubel, D. H. & Wiesel, T. N. *J. Physiol., Lond.* **160,** 106–154 (1962).

Morphine and δ-opiate agonists locally stimulate *in vivo* dopamine release in cat caudate nucleus

M. F. Chesselet, A. Chéramy, T. D. Reisine & J. Glowinski

Groupe NB, INSERM U.114, Collège de France, 11, place Marcelin Berthelot, 75231 Paris Cedex 5, France

The peripheral administration of morphine[1-4] or the intraventricular injection of either Met-enkephalin or D-Ala^2-Met-enkephalinamide[4] (a synthetic enkephalin analogue resistant to enzyme degradation) enhanced dopamine (DA) turnover in the rat striatum. These effects, in turn, could be blocked by the opiate antagonist naloxone[1,4]. As the striatum contains a high density of opiate receptors[5], these drugs may be affecting DA metabolism by a local action. This hypothesis is supported by the evidence that morphine still enhanced striatal DA turnover shortly after transection of the nigro-striatal DA pathway[6]. Furthermore, D-Ala^2-Met-enkephalinamide directly injected into the striatum significantly increased striatal DA turnover[4]. However, the mechanism by which opiates stimulate DA turnover is not clear because they have not yet been shown to enhance DA outflow in *in vitro* release studies[7-9]. We have therefore now examined *in vivo* the local effects of opiates on the release of DA in the cat caudate nucleus. We find that both morphine and more markedly, D-Ala^2-Met-enkephalinamide stimulate DA release. The action of morphine is antagonized by naloxone, but that of the synthetic enkephalin is not. Because the effects of D-Ala^2-Met-enkephalinamide are similar to those of the peripheral δ-receptor agonist Tyr-D-Ser-Gly-Phe-Leu-Thr, we conclude that the synthetic enkephalin acts on δ-receptors.

Experiments were carried out on adult cats of both sexes lightly anaesthesized with halothane, implanted with a push–pull cannula in one caudate nucleus to measure continuously the release of ^{3}H-DA synthesized from ^{3}H-tyrosine as previously described[10]. In these conditions, the spontaneous release of ^{3}H-DA is dependent on nerve activity and can be influenced by the local application of various transmitters[11]. The technique consists in the delivery (1.8 ml h^{-1}) of an artificial cerebrospinal fluid containing ^{3}H-tyrosine (50 Ci mmol^{-1}, 50 μCi ml^{-1}) to the push–pull cannula and in the estimation of ^{3}H-DA in serial

10-min superfusate fractions. ^{3}H-DA is separated from ^{3}H-tyrosine and ^{3}H-labelled metabolites by ion-exchange chromatography and alumina adsorption[12]. The drugs tested were added for 30 min in the superfusing fluid 160 min after the beginning of the superfusion with ^{3}H-tyrosine, that is, at least 1 h after the steady-state level of the spontaneous ^{3}H-DA release had been reached. As the effect of a single application of a compound can only be estimated in one experiment with this preparation and several experiments are required for statistical analysis, only one concentration of each compound was used in most cases.

Morphine (10^{-6} M) significantly stimulated the local release of ^{3}H-DA (Fig. 1*a*). This effect was seen during the last 20 min of the drug application and persisted in the two following fractions after its removal from the superfusing fluid. Naloxone (10^{-6} M) simultaneously added with morphine (10^{-6} M) to the superfusing fluid completely prevented the stimulatory effect of morphine on ^{3}H-DA release (Fig. 1*b*). In fact, the pattern of ^{3}H-DA release seen in the presence of both drugs was similar to that observed with naloxone (10^{-6} M) alone, which induced a delayed reduction in ^{3}H-DA release (Fig. 1*c*).

D-Ala^2-Met-enkephalinamide, used at a concentration (10^{-6} M) similar to that of morphine also stimulated ^{3}H-DA release (Fig. 2*a*), but in this case, the stimulatory effect was immediate, more pronounced and biphasic. The initial increase in ^{3}H-DA release lasted only 10 min whereas the second increase in ^{3}H-transmitter release was seen as soon as the enkephalin analogue was removed from the superfusing medium. In contrast to the results observed with morphine, naloxone (10^{-6} M) did not significantly modify the biphasic stimulatory effect of D-Ala^2-Met-enkephalinamide (10^{-6} M) when simulataneously added with the agonist (Fig. 2*b*). It was still ineffective at a concentration of 10^{-5} M (data not shown). Higher concentrations of naloxone were not used because they have been shown to exert non-specific effects[13,14].

Pharmacological studies on peripherial tissues have revealed the presence of various types of opiate receptor[15]. Furthermore, recent binding studies with striatal membranes using ^{3}H-morphine and ^{3}H-D-Ala^2-D-Leu^5-enkephalin as ligands have revealed a twofold higher amount of δ versus μ opiate receptors in the rat striatum[16]. As morphine seems less effective than the enkephalin derivative in stimulating the *in vivo* release of ^{3}H-DA and as the effect of D-Ala^2-Met-enkephalinamide was not antagonized by an equimolar concentration of naloxone, the effects of this derivative in our experiments could be mainly mediated by δ-receptors. Indeed, although naloxone can block various peripheral opiate receptors, higher concentrations of this antagonist are required to block δ- than μ-receptors[15]. Moreover, naloxone was shown to be 20 times less potent on δ- than on μ-receptors in binding studies with rat striatal membranes[16].

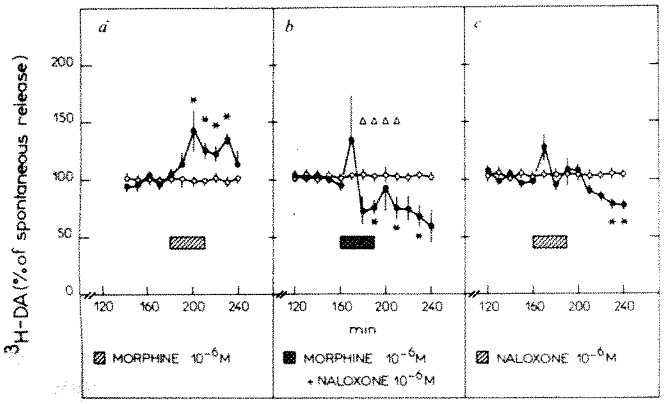


Fig. 1 Effects of morphine (*a*), combined morphine and naloxone (*b*) or naloxone (*c*) applications into the cat caudate nucleus on the release of ^{3}H-dopamine (^{3}H-DA). A push–pull cannula was implanted in the caudate nucleus in halothane-anaesthetized cats. This structure was superfused with an artificial cerebrospinal fluid containing L-[3, 5-^{3}H]tyrosine (50 Ci mmol^{-1}, 50 μCi ml^{-1}, 1.8 ml h^{-1}). ^{3}H-DA was estimated in 10-min successive superfusate fractions. An average of 0.6 nCi of ^{3}H-DA was released in superfusate fractions (30 times the blank value). Morphine (10^{-6} M), naloxone (10^{-6} M) or both morphine (10^{-6} M) and naloxone (10^{-6} M) were introduced for 30 min into the superfusing fluid (hatched bars). In each animal, ^{3}H-DA in each successive fraction was expressed as a percentage of an average spontaneous release calculated from the five fractions collected before the treatment. Data are the mean ± s.e.m. of results obtained with 6(*a*), 5 (*b*) or 14 (*c*) animals (●). *$P < 0.05$ when compared with corresponding control values obtained in six untreated animals (○). $\Delta P < 0.05$ when compared with values obtained in the presence of morphine (10^{-6} M) alone.

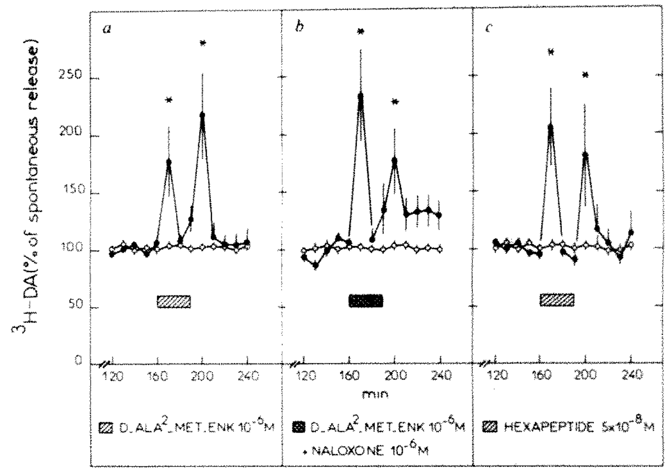


Fig. 2 Effects of D-Ala^2-Met-enkephalinamide (D-Ala^2-Met-Enk) in the absence (*a*) or presence (*b*) of naloxone and of Tyr-D-Ser-Gly-Phe-Leu-Thr (*c*). Seven (*a*), height (*b*) or nine (*c*) animals were treated as described in Fig. 1 legend except that D-Ala^2-Met-enkephalinamide (10^{-6} M) or D-Ala^2-Met-enkephalinamide (10^{-6} M) and naloxone (10^{-6} M) or the hexapeptide (5×10^{-8} M) were applied to the caudate nucleus for 30 min (hatched bars). Data are calculated and expressed as in Fig. 1. * $P < 0.05$ when compared with control values obtained in six untreated cats (○).

Gacel and his collaborators have recently synthesized an hexapeptide, Tyr-D-Ser-Gly-Phe-Leu-Thr, which is 620 times more potent on mouse vas deferens δ-receptors than on guinea pig ileum μ-receptors[17], and so we tested this compound in our preparation. At a concentration (5×10^{-8} M) 100 times lower than the ED_{50} value for its action on peripheral μ-receptors, Tyr-D-Ser-Gly-Phe-Leu-Thr had a marked biphasic stimulatory effect on ^{3}H-DA release similar to that observed with D-Ala^2-Met-enkephalinamide (10^{-6} M) (Fig. 2*c*), further suggesting a δ-receptor-mediated action of D-Ala^2-Met-enkephalinamide on ^{3}H-DA release.

These experiments demonstrate for the first time that opiates stimulate the *in vivo* release of DA in the caudate nucleus, and support previous proposals that morphine or enkephalin derivatives modulate DA metabolism by a local action in the striatum[4]. In fact, as morphine, D-Ala^2-Met-enkephalinamide and the agonist Tyr-D-Ser-Gly-Phe-Leu-Thr were directly applied into the caudate nucleus in our experiments, the opiate receptors involved in the effects observed must be located within this structure. A stimulation of the *in vivo* spontaneous release of ^{3}H-DA in the cat caudate nucleus has also been seen during the local application of acetylcholine (10^{-5} M)[18] or γ-aminobutyric acid (10^{-5} M)[19], which induced a biphasic stimulatory effect similar to that observed with D-Ala^2-Met-enkephalinamide and Tyr-D-Ser-Gly-Phe-Leu-Thr. The rapid interruption of the initial stimulatory effect seen in all these cases could be attributed to the action of the released DA on DA autoreceptors[20].

Combined lesion and binding studies in the rat have revealed that opiate receptors are heterogeneously distributed within the striatum. Some are located on DA terminals[5] while others are present on striatal cell bodies or dendrites[21]. Furthermore, as opiate binding sites are still detectable after the destruction of the nigrostriatal DA neurones by 6-hydroxydopamine and the kainic acid-induced degeneration of striatal perikaraya, these remaining binding sites could be located on striatal afferent fibres other than the DA terminals[21]. Opiate receptors located on striatal cell bodies and dendrites do not seem to be essential for the regulation of DA metabolism because opiates enhance DA turnover after kainic acid-induced degeneration of the striatal perikarya[4]. The opiate receptors located on DA terminals or on other afferent fibres could be involved in the changes in DA release observed in our experiments.

In vitro studies have shown that several transmitters exert a direct or indirect presynaptic influence on DA release[11]. In preliminary experiments, we examined the effects of opiates on the spontaneous release of ^{3}H-DA synthesized from ^{3}H-tyrosine in rat striatal slices and found that morphine (10^{-5} M, 5×10^{-6} M) was ineffective whereas D-Ala^{2}-Met-enkephalin-amide (10^{-6} M) significantly increased ^{3}H-DA release. This further suggests that morphine and D-Ala^{2}-Met-enkephalinamide by different processes.

Opiates have been shown to modulate the *in vitro* release of acetylcholine in the myenteric plexus[22], substance P in the trigeminal nuclei[23] and noradrenaline in the cerebral cortex[8]. In all these studies the effects of opiates were antagonized by naloxone, as was the case in our *in vivo* experiments with the morphine-induced release of ^{3}H-DA. Further studies will be required using pure μ-receptor agonists to determine whether or not the action of morphine is mediated by μ-receptors. In contrast, naloxone (10^{-6} M and 10^{-5} M) did not prevent the effect of D-Ala^{2}Met-enkephalinamide (10^{-6} M) on ^{3}H-DA release, suggesting that μ opiate receptors were not involved.

Biochemical and immunohistochemical studies have demonstrated a rich innervation of the striatum by enkephalinergic neurones[24,25], and experiments in animals implanted with push–pull cannulae have shown that Met-enkephalin is spontaneously released *in vivo* from the cat caudate nucleus (3.5 pg per fraction per 0.5 ml)[25]. It thus remains to be established if endogenous opiates regulate the release of DA from DA terminals. The availability of a specific and potent antagonist of opiate receptors of the δ-type will be required to determine the precise contribution of enkephalinergic neurones in the presynaptic regulation of DA release in the caudate nucleus.

D-Ala^{2}-Met-enkephalinamide was provided by Dr Y. Audigier (Toulouse) and the δ-agonist hexapeptide by Dr Roques (Paris). We thank them for their help and fruitful discussions and Mr M. Desban for assistance. Naloxone was given by Endo Laboratories. This study was supported by grants from INSERM (ATP 58.78.90), DRET (78.004), and Rhône-Poulenc SA. T. D. Reisine is a NINCDS fellow (1 F32 NS 0 6467.01) and A. C. is a research fellow of Rhône-Poulenc SA.

Received 22 September 1980; accepted 9 March 1981.

1. Kuschinsky, K. & Hornykiewicz, O. *Eur. J. Pharmac.* **19,** 119–122 (1972).
2. Gauchy, C. *et al. Eur. J. Pharmac.* **22,** 311–319 (1973.
3. Westerink, B. H. C. & Korf, J. *Eur. J. Pharmac.* **38,** 281–291 (1976).
4. Biggio, G. *et al. Science* **200,** 552–554 (1978).
5. Pollard, H., Llorens-Cortes, C. & Schwartz, J. C. *Nature* **268,** 745–747 (1977).
6. Mcleman, P. & Bruinvels, J. *Nature* **281,** 686–687 (1979).
7. Loh, H. H. *et al. Nature* **264,** 567–568 (1976).
8. Arbilla, S. & Langer, S. Z. *Nature* **271,** 559–560 (1978).
9. Subramanian, N. *et al. Naunyn Schmiedebergs Archs Pharmak.* **299,** 163–166 (1977).
10. Nieoullon, A. *et al. J. Neurochem.* **28,** 819–828 (1977).
11. Glowinski, J. *et al. The Neurobiology of Dopamine* (eds Horn, A. Korf, J. & Westerink, B. 199–216 (Academic, New York, 1979).
12. Nieoullon, A. *et al. Brain Res.* **145,** 69–83 (1978).
13. Sawynok, J. *et al. Life Sci.* **25,** 1621–1632 (1979).
14. Gruol, D. L. *et al. Brain Res.* **198,** 323–332 (1980).
15. Lord, J. A. H. *et al. Nature* **267,** 495–499 (1977).
16. Fields, H. L. *et al. Nature* **284,** 351–353 (1980).
17. Gacel, G., Fournie-Zaluski, M. C. & Roques, B. *FEBS Lett.* 245–247 (1980).
18. Giorguieff-Chesselet, M. F. *et al. Neurosci. Lett.* **14,** 177–182 (1979).
19. Chéramy, A. *et al. Eur. J. Pharmac,* **48,** 281–295 (1978).
20. Westfall, T. C. *et al. Naunyn Schmiedebergs Archs Pharmak.* **292,** 279–287 (1976).
21. Pollard, H. *et al. Brain Res.* **151,** 392–398 (1978).
22. Szerb, J. *Naunyn Schmiedebergs Archs Pharmak.* **311,** 119–127 (1980).
23. Jessel, T. & Iversen, L. L. *Nature* **268,** 549–551 (1977).
24. Sar, M. *et al. J. comp. Neurol.* **182,** 17–38 (1978).
25. Cesselin, F. *et al. Neuroscience* (in the press).

Heavy meromyosin cross-links thin filaments in striated muscle myofibrils

Julian Borejdo & Avraham Oplatka

Polymer Department, The Weizmann Institute of Science, Rehovot, Israel 76100

Studies of the interaction of proteolytic fragments of myosin with actin filaments in solution have indicated that the two heads of the myosin molecule[1] can bind to adjacent subunits in the same actin filament[2,3]. However, geometrical considerations led Offer and Elliott[4] to the conclusion that in the highly ordered filament lattice present in insect flight muscle, tighter binding would occur if the two heads bound to different actin filaments. Indeed, rigidity measurements[5], electron microscope observations[6] and superprecipitation experiments[7] support the idea that heavy meromyosin (HMM), the double-headed myosin fragment, can cross-link actin filaments in solution. We now provide an experimental demonstration of the cross-linking of thin filaments in muscle by double-headed myosin species. In striated muscle, the thin filament lattice can be stabilized by either the Z-bands or complex formation with myosin. Removal of the Z-bands by a proteolytic enzyme[8] or extraction of myosin, leaves the actin-containing I-bands intact. However, if both Z-bands and myosin are removed, the I-bands should disintegrate into single, dispersed, filaments. We found that this could be prevented by preincubation with HMM, but the single-headed HMM subfragment-1 (HMM S-1) was totally ineffective. This strongly suggests that the stabilization by HMM of the filamentous lattice is a result of inter-filament cross-linking.

Myofibrils were prepared from glycerinated rabbit psoas muscle, homogenized in the Omni-Mixer (see Fig. 1 legend for details). HMM and HMM S-1 were prepared by the chymotryptic digestion of myosin described elsewhere[9] and HMM was further purified by fractional precipitation with ammonium sulphate. Myosin contamination of the fragments was assessed by SDS-polyacrylamide gel electrophoresis; the HMM S-1 preparations had no detectable contamination and that in HMM was estimated to be $<1\%$ (by weight).

Figure 1 shows a series of manipulations of a single myofibril in rigor solution. An intact myofibril (Fig. 1*a*) was treated with the myosin-extracting solution, resulting in a myosin-free ghost (Fig. 1*b*) in which the Z-lines are clearly visible. If such ghosts are then treated with a trypsin solution, they disintegrate within a few seconds (not shown). The ghosts were first incubated with HMM, which when bound to the thin filaments enhanced the phase image of the ghost, but the Z-lines were still clearly visible (Fig. 1*c*). After treatment with trypsin-containing solution, the Z-lines were digested and disappeared completely in 60 s (Fig. 1*d*). Nevertheless, as clearly demonstrated in Fig. 1*d*, and reproduced in 12 identical experiments with 3 different HMM preparations (including HMM containing 'short' S-2), the ghost remained intact. If an identical procedure is followed with HMM S-1 instead of HMM, the ghosts disappear within a few seconds after the application of the Z-line-digesting solution (not shown).

We considered the possibility that residual myosin in the HMM preparations (at most 1% by weight, see above) was responsible for this effect. When the myofibrilar ghosts were pre-incubated with a rigor solution containing 1.5 mg ml^{-1} HMM S-1 and 0.15 mg ml^{-1} myosin, the removal of the Z-lines led to a rapid disappearance of the ghosts. Similarly, pre-incubation of the ghosts with the same ratio of HMM S-1 to myosin in a high ionic strength solution (to ensure full penetration of myosin into the thin filament network), followed by washing with rigor solution and digestion of the Z-lines, results

0028-0836/81/220322—02$01.00

in a rapid (within a few seconds) disappearance of the ghosts. It is also conceivable that the digestion of the Z-lines by trypsin in the presence of HMM is much slower than that in the presence of HMM S-1 (the larger HMM could, for instance, sterically block access of trypsin to the Z-lines). That this could not be the reason for the stability of ghosts pre-incubated with HMM was demonstrated by the fact that the addition of as little as 1 μM MgATP (in the presence of 4 mM phosphocreatine and 0.4 mg ml^{-1} creatine kinase) to such ghosts after treatment with trypsin led to their instantaneous disappearance. These control experiments clearly show that it is not the contaminating myosin in the HMM preparation or the incomplete removal of the Z-lines that is responsible for the maintenance of the structural integrity of the ghosts.

It is similarly possible to demonstrate the cross-linking ability of the indigenous myofibrilar myosin. Perfusion of the myofibril with a solution containing 0.6 M KCl, 5 mM Na-phosphate buffer, *p*H 7.0, causes the thick filaments to dissociate into presumably individual myosin molecules. Z-lines are then removed by the above solution containing, in addition, 0.5 mg ml^{-1} trypsin. Myofibrils thus treated still maintain structural integrity.

Although these experiments do not elucidate which fraction of HMM molecules is involved in the inter-filament cross-linking or whether this type of binding is responsible for the inability of HMM fully to saturate actin binding sites in myofibrilar ghosts[10], they clearly demonstrate the ability of the double-headed myosin moiety to bind different actin filaments *in vivo*. Future research should examine the possibility, which is now more likely, that a similar two-filament interaction occurs during muscle contraction. Whether or not the two heads are identical, any discussion of the mechanism of contraction will have to take into consideration the capability of a myosin molecule to interact with two different actin filaments.

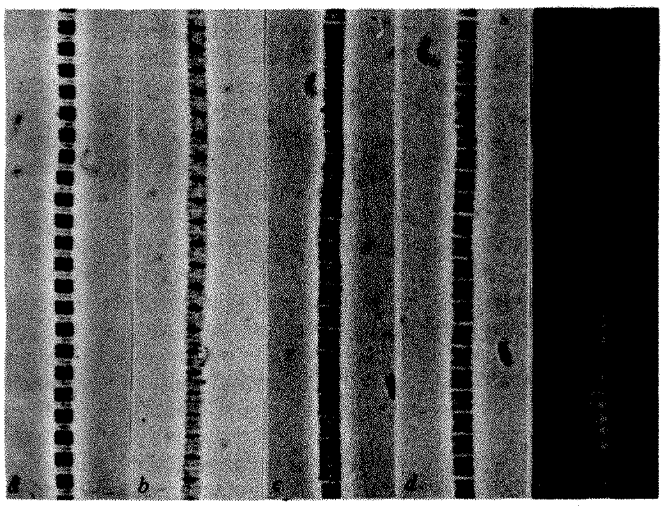


Fig. 1 Myofibrils were obtained by homogenizing glycerinated fibres in rigor solution (in a Sorvall Omni-Mixer for 6 s at speed setting 7, in ice). The myofibrils were filtered through gauze and suspended in a rigor solution containing 0.1 mM sodium azide and 1 mg ml^{-1} phenylmethylsulphonylfluoride. The myofibril bathing solutions were changed by applying a drop of the appropriate solution at one end of the coverslip and soaking it up with blotting paper at the opposite end. All experiments were done at room temperature using the Zeiss Axiomat microscope, ×100 Planapochromat objective, N.A. = 1.3 The phase retardation of the diffraction ring was $-\pi/2$ and its amplitude transmission 0.1. *a*, Intact myofibril in rigor solution; *b*, the same myofibril after myosin extraction with the modified Hasselbach–Schneider solution containing 0.5 M KCl, 0.1 M Na-phosphate buffer, *p*H 7.0, 5 mM $MgCl_2$ and 10 mM Na-pyrophosphate. This solution was freshly prepared before each experiment. *c*, The same myofibril after perfusion with a rigor solution. *d*, The same myofibril after removal of the Z-line using rigor solution containing 0.5 mg ml^{-1} trypsin, and a thorough wash with the rigor solution. *e*, Fluorescence image of *d*. The HMM in this experiment was labelled with 2 mol iodoacetamido-flourescein per mol HMM by the method described elsewhere[10].

We thank Rina Levy for technical assistance and the Muscular Dystrophy Association and the US–Israel Binational Foundation for financial support. J. B. is an Established Investigator of the American Heart Association.

Received 4 March; accepted 23 March 1981.

1. Slayter, H. S. & Lowey, S. *Proc. natn. Acad. Sci. U.S.A.* **58,** 1611–1618 (1967).
2. Moore, P. B., Huxley, H. W. & De Rosier, D. J. *J. molec. Biol.* **50,** 279–295 (1970).
3. Craig, R. C. *et al. J. molec. Biol.* **140,** 35–55 (1980).
4. Offer, G. & Elliot, A. *Nature* **271,** 325–329 (1978).
5. Abe, S. I. & Maruyama, K. *Biochim. biophys. Acta* **160,** 160–172 (1974).
6. Trinick, J. & Offer, G. *J. molec. Biol.* **133,** 549–556 (1979).
7. Borejdo, J., Garty, N., Shimon, T., Himmelhoch, S. & Oplatka, A. *FEBS Lett.* **105,** 101–104 (1979).
8. Goll, D. E. *et al. J. Anim. Sci.* **33,** 963–982 (1971).
9. Weeds, A. G. & Pope, B. *J. molec. Biol.* **111,** 129–157 (1977).
10. Borejdo, J. & Assulin, O. *Biochemistry* **19,** 4913–4921 (1980).

Thromboxane and prostacyclin release from ischaemic myocardium in relation to arrhythmias

Susan J. Coker, J. R. Parratt, I. McA. Ledingham & I. J. Zeitlin

Departments of Physiology and Pharmacology,
University of Strathclyde, Glasgow G1 1XW, and
Department of Surgery, University of Glasgow (Western Infirmary),
Glasgow G11 6NT, UK

There has been much interest in the possible physiological and pathophysiological roles of myocardial prostaglandins[1]. One particularly intriguing suggestion[2] is that they may be capable of protecting the heart against arrhythmogenic stimuli. There is certainly no doubt that some prostaglandins can protect against a variety of arrhythmias induced either chemically or by coronary artery ligation[2–4]. For example, PGE_2 and prostacyclin are effective against the early post-infarction arrhythmias that result from acute coronary artery ligation in anaesthetized dogs[3] and rats[4]. In a previous study[5,6] we could find no evidence that PGE_2 (or $PGF_{2\alpha}$) was released from the canine myocardium at a time when ventricular ectopic activity was pronounced. We now demonstrate that, in contrast, both thromboxane and prostacyclin are released from the acutely ischaemic myocardium and that the balance between thromboxane and prostacyclin release in the period immediately following coronary artery ligation is related to the occurrence of arrhythmias.

The studies were performed in greyhounds (21–34 kg) anaesthetized with chloralose (80 mg per kg, intravenously) following induction with sodium thiopentone (25 mg per kg, intravenously). The animals were respired with oxygen using a positive pressure ventilation pump and catheters were placed in the aorta, coronary sinus, left ventricle and pulmonary artery under fluoroscopic control. They were used for blood sampling, pressure measurement and for the determination of cardiac output by thermodilution. Following a left thoracotomy the heart was suspended in a pericardial cradle and a ligature placed around the left anterior descending coronary artery. The vein adjacent to this artery was catheterized using the Seldinger technique such that the tip lay (after coronary artery ligation) within the ischaemic region. Blood samples were taken from the various sites 5 min before acute coronary artery ligation and at 2, 10 and 30 min afterwards. They were analysed for blood gases and *p*H[7] and also for thromboxane B_2 and 6-keto $PGF_{1\alpha}$ (the major breakdown products of thromboxane A_2 and prostacyclin, respectively), using radioimmunoassay techniques developed by us[6,8], based on the methods of Dray *et al.*[9].

Blood samples for thromboxane and prostacyclin assays were immediately placed in plastic tubes containing indomethacin and EDTA and kept on ice until centrifugation (not more than

0028–0836/81/220323—02$01.00

40 min later) at 2,000g for 10 min. The plasma was removed and stored at −20 °C until assay. After acidification, aliquots of plasma were extracted with redistilled ethyl acetate and these extracts were then evaporated to dryness under reduced pressure at 37 °C. The recovery of thromboxane B_2 or 6-keto $PGF_{1\alpha}$ was monitored using tritiated internal standards. Sample extracts and standards were then redissolved in phosphate-buffered saline and tritiated thromboxane B_2 or 6-keto $PGF_{1\alpha}$ (NEN) was added followed by the appropriate specific antibody (Pasteur Institute). After overnight incubation at 4 °C, the antibody-bound and free prostaglandins were separated using dextran-coated charcoal. An aliquot of the antibody-bound fraction was placed in Biofluor scintillant and counted in a Packard Tri-Carb 460 liquid scintillation counter. With the above procedures the detection limits were 15 and 40 pg for thromboxane B_2 and 6-keto $PGF_{1\alpha}$, respectively.

Acute coronary artery ligation resulted in electrocardiographic and local metabolic changes[7]. These animals exhibited varying numbers of cardiac arrhythmias ranging from 5 to 1,085 ventricular ectopic beats in the first 30 min after ligation. Three animals out of 12 died in ventricular fibrillation during this time.

There was no evidence of myocardial thromboxane release in the control state (arterial, coronary sinus and local coronary venous concentrations of 181 ± 43, 188 ± 50 and 202 ± 38 pg ml^{-1}, respectively). After coronary artery ligation, however, there was an immediate and marked increase in the thromboxane B_2 concentration in local coronary venous blood (Fig. 1a), evident 2 min after ligation and peaking at 10 min; by 30 min thromboxane values were returning towards control. After ligation there were no significant changes in the concentration of thromboxane B_2 in arterial or coronary sinus blood, which drains the essentially normal myocardium (Fig. 1a). These results indicate, therefore, that coronary artery ligation results in the early release of thromboxane only from the ischaemic region.

Despite marked variations in 6-keto $PGF_{1\alpha}$ concentrations between individual animals (mean arterial concentration before ligation, 462 ± 166 pg ml^{-1}), there was evidence that 6-keto $PGF_{1\alpha}$ was released from the myocardium in the control state, as the coronary sinus concentrations were higher than the arterial values in each animal, the difference ranging from 133 to 364 pg ml^{-1}, mean ± s.e.m. 242 ± 28, $n = 9$. After coronary artery ligation there was a further myocardial release of 6-keto $PGF_{1\alpha}$ but, in contrast to thromboxane release, this occurred rather more gradually and from both the ischaemic and normal regions of the left ventricular wall (Fig. 1b).

Close examination of the results in each individual animal revealed an interesting pattern. At 2 min post-ligation, the animals which exhibited the most ventricular ectopic activity also showed the largest increases in local coronary venous thromboxane B_2 concentrations. This point is illustrated in Fig. 2, which shows that there was a significant positive correlation ($r = 0.863$) between the change in coronary venous thromboxane B_2 at 2 min post-ligation and the number of arrhythmias which had occurred by this time. Conversely, there was a significant negative correlation ($r = 0.912$) between the change in 6-keto $PGF_{1\alpha}$ concentrations and arrhythmias. This suggests that the number of arrhythmias which occur in this period may be related to the balance between thromboxane and prostacyclin release, a balance in favour of thromboxane release being associated with a greater number of arrhythmias.

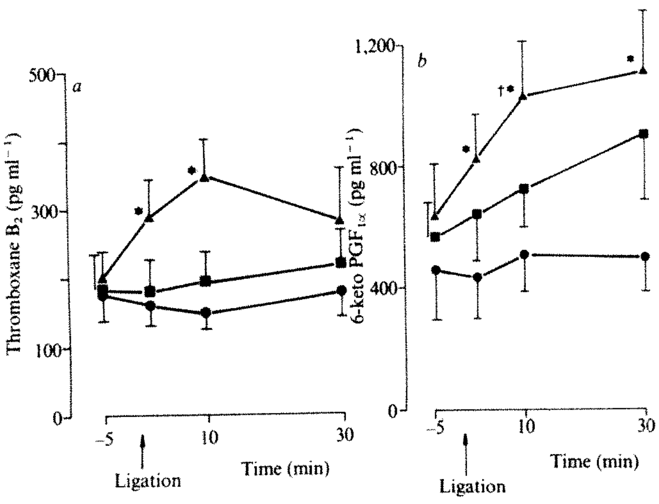


Fig. 1 Thromboxane B_2 (a) and 6-keto $PGF_{1\alpha}$ (b) values in the aorta (●), coronary sinus (■) and local coronary vein (▲) before and 2, 10 and 30 min after coronary artery ligation. Each value represents the mean ± s.e.m., $n = 9$. $^*P < 0.05$ (venous vs arterial), $^\dagger P < 0.05$ (venous vs sinus), dependent t-test.

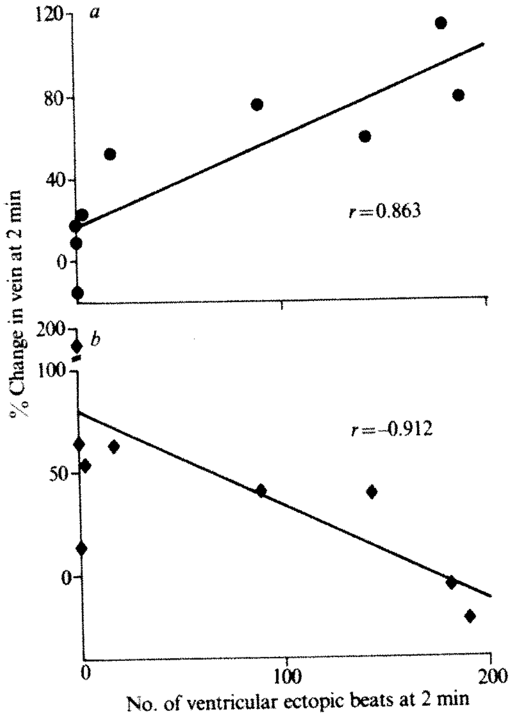


Fig. 2 Correlation between the % change in thromboxane B_2 in the local coronary vein 2 min post-ligation and the number of arrhythmias which had occurred by that time (a) and similarly between the % change in 6-keto $PGF_{1\alpha}$ and arrhythmias (b).

The most likely explanation for this observed relationship between the release of these substances and the severity of early post-infarction arrhythmias concerns their differential effects on the coronary microcirculation. Thromboxane may reduce oxygen delivery to already hypoxic myocardial cells by a combination of local vasoconstriction[10] and mechanical obstruction resulting from platelet aggregation. Prostacyclin, on the other hand, would both inhibit platelet aggregation and dilate small coronary vessels. It is not known whether thromboxane A_2 has direct arrhythmogenic properties but prostacyclin has been shown to have anti-arrhythmic activity in the dog[4]. These results suggest that a stable prostacyclin analogue and/or a specific thromboxane synthesis inhibitor might prove beneficial in the early stages of acute myocardial infarction.

Supported by the Scottish Hospitals Research Trust.

Received 8 January; accepted 19 March 1981.

1. Sivakoff, M., Pure, E., Hseuh, W. & Needleman, P. *Fedn Proc.* **38,** 78–82 (1979).
2. Förster, W. *Acta biol. med. germ.* **35,** 1101–1112 (1976).
3. Harvie, C.J., Collins, G. A., Miyagishima, R. T. & Walker, M. J. A. *Prostaglandins* **16,** 885–900 (1978).
4. Au, T. L. S., Collins, G. A., Harvie, C. J. & Walker, M. J. A. *Prostaglandins* **18,** 707–720 (1979).
5. Coker, S. J., Marshall, R. J., Parratt, J. R., Siddiqui, H. H. & Zeitlin, I. J. *J. molec. cell. Cardiol.* **11,** Suppl 11, 8 (1979).
6. Coker, S. J., Marshall, R. J., Parratt, J. R. & Zeitlin, I. J. *J. molec. cell. Cardiol.* (in the press).
7. Marshall, R. J., Parratt, J. R. & Ledingham, I. McA. *Cardiovascular Res.* **8,** 204–215 (1974).
8. Coker, S. J. thesis, Univ. Strathclyde (1980).
9. Dray, F., Charbonnel, B. & Maclouf, J. *Eur. J. clin. Invest.* **5,** 311–318 (1975).
10. Terashita, Z-I., Fukui, H., Nishikawa, K., Hirata, M. & Kikuchi, S. *Eur. J. Pharmac.* **53,** 49–56 (1978).

Sodium current depression by lidocaine and quinidine in isolated ventricular cells

K. S. Lee, J. R. Hume, W. Giles* & A. M. Brown

Department of Physiology and Biophysics, University of Texas
Medical Branch, Galveston, Texas 77550, USA

A number of local anaesthetics are now widely used as cardiac antiarrhythmic agents, but the cellular basis for these actions in the heart remains controversial[1]. One hypothesis is that agents such as lidocaine and quinidine selectively depress the fast inward current, I_{Na}. We have now tested this hypothesis using recently developed techniques for isolating and voltage clamping rat ventricular cells. Our results provide direct evidence that therapeutic doses (10–20 μM) of lidocaine or quinidine strongly inhibit I_{Na}, and give some clues about the mechanism of this inhibition.

Weidmann[2] first reported that the local anaesthetic cocaine depresses the maximum upstroke velocity ($\dot{V}_{max}$) of the action potential in sheep Purkinje fibres; his results also suggested that this effect was produced by a voltage-dependent shift in the inactivation mechanism governing I_{Na}. Johnson and MacKinnon[3] and Heistracher[4] later showed that quinidine sulphate depressed $\dot{V}_{max}$ in atrial tissue and papillary muscle. Their results emphasized that the amount of depression of $\dot{V}_{max}$ was strongly dependant on stimulus frequency. A similar frequency-dependent, or use-dependent, depression of $\dot{V}_{max}$ by lidocaine, quinidine and various other local anaesthetic agents in cardiac tissue has now been reported[5,6]. However, $\dot{V}_{max}$ in most experimental tests provides only an indirect and nonlinear measure of the maximum available sodium conductance, $\bar{g}_{Na}$ (refs 7–11).

A reduction in I_{Na} by local anaesthetics has been previously observed in voltage-clamp experiments done in squid axon, frog node of Ranvier, and frog skeletal muscle (see ref. 12). These data show that local anaesthetics (0.1–5 mM at pH 7.2–7.4) produce a large tonic block as well as a use-dependent inhibition of I_{Na}. Most of the data from these tissues are consistent with a model (see ref. 12) of use-dependence in which the cationic drug molecules bind to a site within the sodium channel, and this drug binding is enhanced when the I_{Na} channels are in the inactivated state. There is no direct evidence that this mechanism is applicable to cardiac muscle, therefore we have attempted to study the actions of lidocaine and quinidine on I_{Na} in single cells derived from rat ventricular tissue. These two drugs were chosen because the kinetics of their use-dependent inhibition of I_{Na} has been postulated to be very different[5]. Our initial experiments were restricted to semiquantitative measurements of I_{Na} at normal pH in the presence of therapeutic doses ($\sim 10^{-5}$ M) of these two drugs. Emphasis was placed on possible drug-induced changes in the inactivation of I_{Na}.

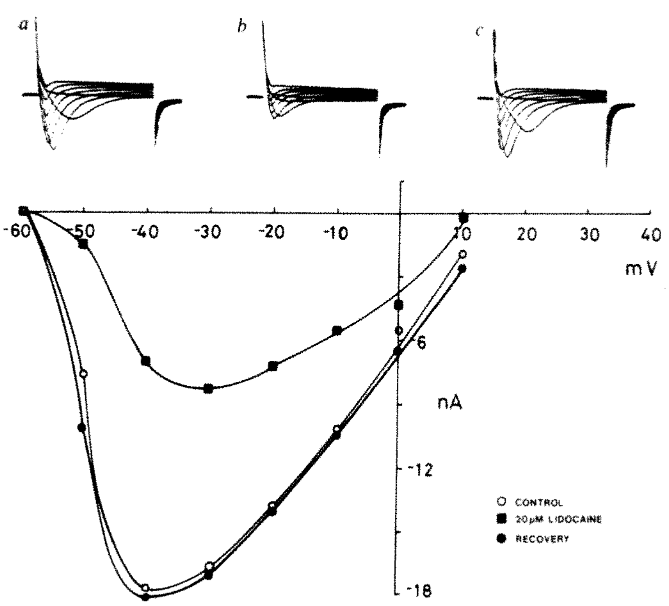


Fig. 1 Effects of lidocaine-HCl (20 μM) on I_{Na} in isolated cells from the rat ventricle. *a–c*, Are superimposed current records resulting from 10-ms depolarizations (+20–+90 mV) applied at 0.1 Hz from a holding potential of −80 mV. *b*, Shows the reduction in I_{Na} after exposure to the drug for 10 min. Note the marked reduction in I_{Na} with relatively little change in steady background current. The bottom part of the figure shows corresponding *I–V* relationships for I_{Na}. Each data point was obtained by subtracting the maximum inward current from the steady background current recorded at the end of the depolarizing pulse. ■, The *I–V* relationship obtained for lidocaine (20 μM); ○ and ● represent control and recovery tests, respectively.

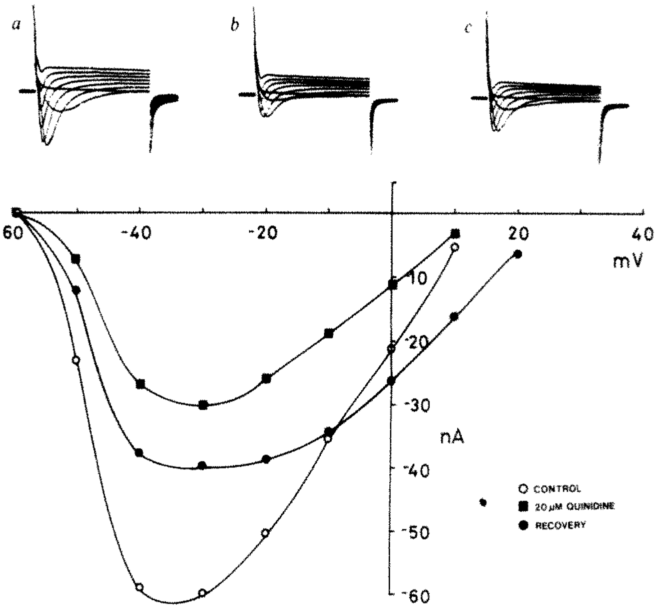


Fig. 2 Effects of quinidine-HCl (20 μM) on I_{Na} in isolated cells from the rat ventricle. For details of experimental protocol, see Fig. 1 legend. *b*, And the *I–V* relationship (■) indicate the voltage-clamp data obtained in the presence of the drug.

Isolated cells from the ventricle of adult rats were prepared for internal perfusion and voltage clamp of these cells, using the procedure of Powell and Twist[13]. A suction pipette technique[14,15] which was developed in this laboratory was utilized. In some experiments, two suction pipettes were used to achieve more effective series resistance compensation. As the resulting clamp data were qualitatively similar to those obtained using a single suction pipette with series resistance compensation, all results were combined. The superfusion solution (pH 7.4) contained 44 mM Na-isethionate, 52.4 mM Cs aspartate, 2.8 mM NgSO$_4$, 1.1 mM MgCl$_2$, 1.1 mM CaCl$_2$, 5.5 mM glucose, 80 mM sucrose and 8 mM Tris base. Thus, $[Na]_o$ was reduced to about one-third normal, and sodium was the only permeant monovalent cation in the external solution. The solution which perfused the fibre internally (pH 7.2) consisted of 135 mM Cs aspartate, 17 mM NaHCO$_3$, 5 mM glucose and 2 mM Tris base. Transmembrane voltage and current waveforms were displayed on a storage oscilloscope and digitized every 2 μs using a Nicolet 1170 digital oscilloscope. Digitized records were stored on tape (Kennedy 9700 digital tape recorder) for off-line analysis. All experiments were done at room temperature.

---

*To whom all correspondence should be addressed.

0028–0836/81/220325—03$01.00

Figure 1 shows records of I_{Na} obtained in the presence (*b*) and absence (*a* and *c*) of 20 μM lidocaine–HCl. A plot of the *I–V* relation for I_{Na} is shown in the bottom part of Fig. 1. In this and all subsequent analyses, I_{Na} was measured as the ΔI between the peak of I_{Na} and the steady level of current at the end of a 10-ms depolarizing voltage clamp pulse. Possible use-dependent depression of I_{Na} was minimized by applying clamp pulses at 0.1 Hz, after the quiescent preparation had been exposed to the drug for 3–5 min. It is apparent from Fig. 1 that lidocaine-HCl strongly reduces I_{Na}, and that this effect is selective and fully reversible after a 10–20-min wash period. Similar results (50–80% tonic or resting reduction in I_{Na}) were obtained in six additional experiments.

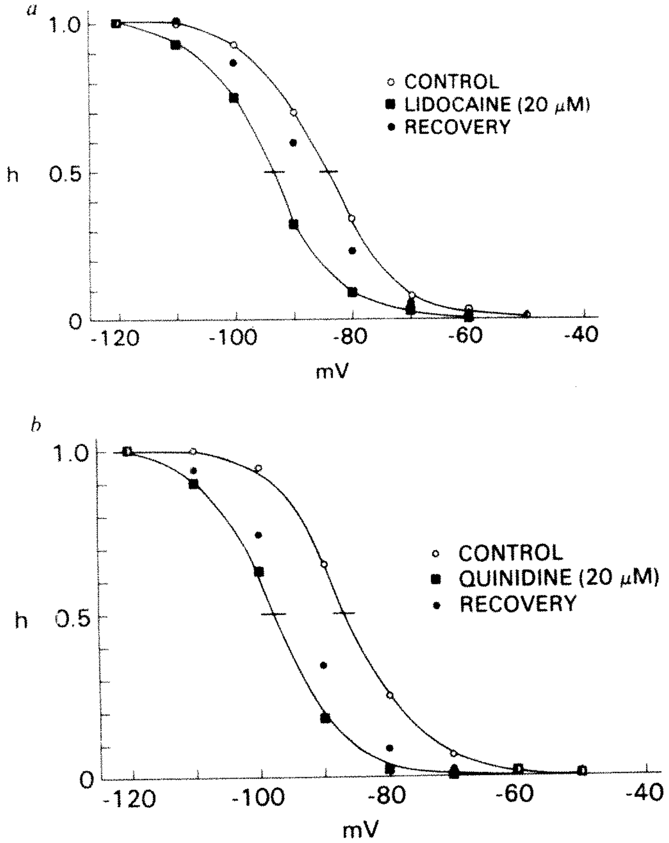


Fig. 3 *a*, Effect of lidocaine-HCL (20 μM) on steady-state inactivation of I_{Na}. Voltage dependence of the inactivation process was measured by applying a series of paired pulses at 0.1 Hz. A 500-ms prepulse of variable magnitude (−140 to −50 mV) was followed immediately by a standard test pulse of +40 mV for 10 ms. Points on the curves (○, control; ■, lidocaine-HCl and ●, recovery) represent the normalized peak inward currents corresponding to each prepulse voltage level. The horizontal lines represent 50% inactivation for the control and drug curves. *b*, Effect of quinidine-HCl (20 μM) on steady-state inactivation of I_{Na}. Same protocol as for *a*. Note that quinidine also produces a voltage-dependent shift in the h curve. Schwarz *et al.*[21] have drawn attention to the complexities of this measurement when testing agents which have a significant use-dependent effect. In their view, the protocols used in these experiments measure the h curve of those Na channels which are not blocked at rest.

Figure 2 shows data from an identical study of the effects of 20 μM quinidine-HCl on I_{Na}. A selective ($\sim$70%) tonic reduction in the peak I_{Na} was observed in each of seven experiments. However, we did not obtain full reversibility before the preparation began to deteriorate ($\sim$20 min).

We next tested whether this decrease in I_{Na} was produced by drug-induced changes in the inactivation variable which controls I_{Na}. Figure 3 shows that both lidocaine–HCl (*a*) and quinidine–HCl (*b*) shifted the steady-state inactivation curve (produced by 500-ms prepulses) to more negative or hyperpolarized potentials. When such inactivation curves are normalized, the mean voltage shift (±s.e.m.) measured at the midpoint of the curve is: lidocaine-HCl (10.4 ± 1 mV, $n = 5$); quinidine–HCl (13.5 ± 1.9 mV, $n = 5$). These results are similar to the data obtained by Weidmann[2]. Chen *et al.*[16], however, found no quinidine-induced voltage dependent shift in I_{Na} inactivation as measured by $\dot{V}_{max}$.

The finding that most antiarrhythmic agents prolong the refractory period in ventricular tissue has been attributed to their ability to prolong the recovery of I_{Na} from inactivation[17]. We therefore studied the recovery of I_{Na} in the presence and absence of lidocaine-HCl and quinidine-HCl. Recovery at −80 mV was assessed by measuring I_{Na} in response to paired 40-mV, 40-ms depolarizing pulses separated by intervals varying from 20 to 500 ms. Each two-pulse sequence was followed by a 10–15-s rest period. For lidocaine-HCl (20 μM) the half time of recovery was prolonged from 87 ± 14 to 125 ± 14 ms ($n = 6$); for quinidine–HCl the half time of recovery was slowed from 42 ± 5 to 70 ± 9 ms ($n = 7$). The half time of the recovery process was chosen as a measure because the I_{Na} reactivation does not follow a simple exponential time course in these rat ventricular cells[18].

The final series of experiments (see Fig. 4) was done to assess the extent to which frequency-dependent or use-dependent inhibition of I_{Na} contributes to the observed reduction in I_{Na}. Trains of 20 depolarizing voltage clamp pulses (−80 to −40 mV for 40 ms) were applied at frequencies ranging from 0.1 to 5 Hz. No measurable use-dependent effects were observed in either drug when the pulsing frequency was less than 0.3 Hz. At 5 Hz a small use-dependent inhibition was observed. Note that in Fig. 4*a* the resting or tonic blockade by lidocaine-HCl is 40% of the control I_{Na} whereas frequency-dependent blockade produced an additional 10% inhibition. Similarly, Fig. 4*b* shows that quinidine produces 60% tonic block and that the 5-Hz train produces only an additional 8% inhibition. These results contrast with the previously reported marked modulation of

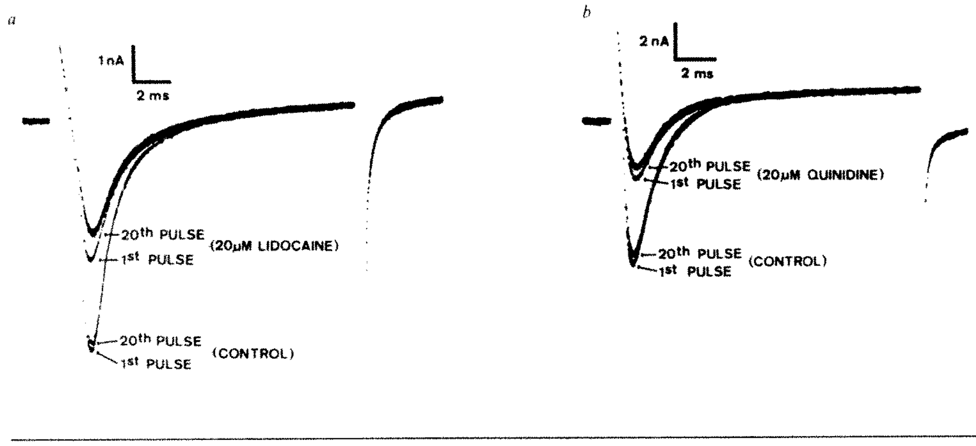


Fig. 4 *a*, Frequency- or use-dependent effects of lidocaine–HCl (20 μM) and *b*, quinidine–HCl (20 μM). Trains of 20 depolarizing voltage clamp pulses (−80 mV to −30 mV for 10 ms) were applied at 5 Hz. The current records in response to pulse no. 1 and pulse no. 20 were superimposed using a storage oscilloscope. When control records has been obtained, the drug was applied to the quiescent preparations, and after 5–10 min the 20 pulses were again applied. Both drugs produced small use-dependent suppression of I_{Na}. Note however, that even in the control test, the I_{Na} in response to pulse 20 was smaller than that elicited by pulse 1.

$\dot{V}_{\max}$ by lidocaine and quinidine[5,6], which implied much larger use-dependent effects on I_{Na} in cardiac muscle. Moreover, they suggest that relatively low doses of two antiarrhythmic agents inhibit I_{Na} in a similar fashion: both produce a large 'resting' or tonic inhibition. The inactivation curve is shifted to the left and an additional use-dependent inhibition is observed although the use-dependent effect is quite small in these experimental conditions.

Before the relative importance of use-dependent compared with tonic blockade of I_{Na} can be assessed conclusively, additional data must be collected in conditions in which the *p*H, the holding potential and the potential to which the cells are depolarized are widely varied. Analogous experiments in squid axon[19] have led to the suggestion that there may be two separate receptors for the resting or tonic block and the frequency-dependent blocking. In addition, the anaesthetic-induced changes in outward current(s) in the heart[20,22] must be studied quantitatively before the basis of these antiarrhythmic actions can be understood fully.

We thank Dr Trevor Powell for demonstrating his technique of rat cell preparation and Ester Lee for technical assistance. This research was supported by a NSF fellowship to J.R.H. and grants from the American Heart Association (77-691 to W.G.) and USPHS (HL-25409 to W.G.; HL-25145 to A.M.B.).

Received 2 October 1980; accepted 12 March 1981.

1. Hauswirth, O. & Singh, B. N. *Pharmac. Rev.* **30,** 5–63 (1979).
2. Weidmann, S. *J. Physiol., Lond.* **129,** 568–582 (1955).
3. Johnson, E. A. & McKinnon, M. G. *J. Pharmac. exp. Ther.* **120,** 460–468 (1957).
4. Heistracher, P. *Naunyn-Schmeidebergs Archs Pharmak.* **269,** 199–212 (1971).
5. Hondeghem, L. & Katzung, B. G. *Circulation* **61,** 1217–1226 (1980).
6. Courtney, K. R. *J. molec. cell. Cardiol.* **12,** 1273–1286 (1980).
7. Schwarz, J. R., Ulbricht, W. & Wagner, H. H. *J. Physiol., Lond.* **233,** 167–194 (1973).
8. Hunter, P. J., McNaughton, P. A. & Noble, D. *Prog. Biophys. molec. Biol.* **30,** 99–144 (1975).
9. Walton, M. & Fozzard, H. A. *Biophys. J.* **25,** 407–420 (1979).
10. Bean, B. P., Cohen, C. J., Colatsky, T. J. & Tsien, R. W. *J. Physiol., Lond.* (in the press).
11. Cohen, I. S. & Strichartz, G. R. *Biophys. J.* **17,** 275–279 (1977).
12. Hille, B. in *Biophysical Aspects of Cardiac Muscle* (eds Morad, M. & Tabatabai, M.) (Academic, New York, 1978).
13. Powell, T. & Twist, V. W. *Biochem. biophys. Res. Commun.* **72,** 327–333 (1976).
14. Lee, K. S., Weeks, T. A., Kao, R. L., Akaike, N. & Brown, A. M. *Nature* **278,** 269–271 (1979).
15. Lee, K. S., Akaike, N. & Brown, A. M. *J. Neurosci. Meth.* **2,** 51–78 (1980).
16. Chen, C., Gettes, L. S. & Katzung, B. G. *Circulation Res.* **37,** 19–29 (1975).
17. Gettes, L. S. & Reuter, H. *J. Physiol., Lond.* **240,** 703–724 (1974).
18. Brown, A. M., Lee, K. S. & Powell, T. *J. Physiol., Lond.* (in the press).
19. Cahalan, M. D. & Almers, W. *Biophys. J.* **27,** 39–56 (1979).
20. Weld, F. M. & Bigger, J. T. *Circulation Res.* **38,** 203–208 (1976).
21. Schwarz, W., Palade, P. T. & Hille, B. *Biophys. J.* **20,** 343–368.
22. Gliklich, J. I. & Hoffman, B. F. *Circulation Res.* **43,** 638–651 (1979).

Calcitonin selectively stimulates 25-hydroxyvitamin D_3-1α-hydroxylase in proximal straight tubule of rat kidney

Hiroyuki Kawashima, Shozo Torikai & Kiyoshi Kurokawa*

Nephrology Section, Medical and Research Services,
Veterans Administration Wadsworth Medical Center, Los Angeles, California 90073, USA
Department of Medicine, the University of California School of Medicine, Los Angeles, California 90024, USA

We have reported previously that the 25-hydroxyvitamin D_3 ($25(OH)D_3$)-1α-hydroxylase (1α(OH)ase) activity was localized exclusively in the proximal convoluted tubules (PCT) in mature vitamin D-deficient rats and that the enzyme activity was largely abolished by parathyroidectomy in the vitamin D-deficient rats with presumed secondary hyperparathyroidism[1]. However Akiba *et al.*[2] found enzyme activity in both PCT and the proximal straight tubules (PST) of the fetal rabbit kidney. Parathyroid hormone and calcitonin, when given *in vivo*, can stimulate the production of 1α, $25(OH)_2D_3$ from $25(OH)D_3$ in vitamin D-deficient rats[3–6], and, moreover, their *in vivo* effects on the 1α(OH)ase in thyroparathyroidectomized vitamin D-deficient rats are additive[6], a finding consistent with different sites of action of these two hormones. In addition, plasma calcitonin levels are elevated in the fetus of several mammalian species[7–10], while they are likely to be low in hypocalcaemic vitamin D-deficient rats. From these observations, we hypothesized that calcitonin could be responsible for stimulating enzyme activity in the PST while parathyroid hormone primarily activates 1α(OH)ase in the PCT. To test this hypothesis, we examined the effect of calcitonin on the 1α(OH)ase activity in defined nephron segments of vitamin D-deficient rats. We report here that the hormone selectively stimulates 1α(OH)ase activity in the PST, whereas this enzyme activity is undetectable in control vitamin D-deficient animals.

Male weanling Holtzman rats were fed a vitamin D-deficient diet containing 0.45% calcium and 0.3% phosphorous for 3–7 weeks[1]. Vitamin D-deficient rats were given graded doses of synthetic salmon calcitonin (Armour) and calcium gluconate, 200 mg per kg subcutaneously every 2 h; the latter was given to prevent a further fall in plasma calcium concentration. The rats were killed 2 hours after the last injection of calcitonin. Control vitamin D-deficient animals received the vehicle for calcitonin (0.9% NaCl) either with or without calcium gluconate. Plasma concentrations of calcium and inorganic phosphate in vitamin D-deficient rats, measured at the time of death, were 6.1 ± 0.4 and 5.8 ± 0.3 mg dl^{-1}, respectively, with no differences between those receiving and not receiving calcium gluconate; calcium and inorganic phosphate values in calcitonin-treated rats were 5.8 ± 0.5 and 5.4 ± 0.7 mg dl^{-1}, respectively. The conversion of ^{3}H-$25(OH)D_3$ to ^{3}H-1α, $25(OH)_2D_3$ was determined in each group of defined tubular segments, microdissected from collagenase-treated kidney slices and incubated in 20 μl Hanks' solution containing 5 mM pyruvate, 8.3 mM glucose, 10 mM HEPES, *p*H 7.4, and 0.5 μM ^{3}H-$25(OH)D_3$ (see Fig. 1). Each incubation contained the specific segments from one animal, and only one incubation was prepared for any given segments from each amimal. The incubations, carried out at 37 °C for 60 min, were terminated by the addition of 50 μl chloroform/methanol (1:2, v/v) followed by lipid extraction[11].

Lipid extracts were subjected to the silica gel 60 TLC; the radioactive peaks were re-extracted and underwent HPLC to identify the metabolites of ^{3}H-$25(OH)D_3$ (ref. 1). The PCT of control rats and both the PCT and PST of calcitonin-treated rats produced a single radioactive peak from ^{3}H-$25(OH)D_3$; this peak co-migrated with authentic 1α, $25(OH)_2D_3$ in both TLC and HPLC. The amounts of 1α, $25(OH)_2D_3$ produced, expressed in fmol ^{3}H–1α, $25(OH)_2D_3$ produced per h per mm

Table 1 1α(OH)ase activity along the nephron of vitamin D-deficient rats given calcitonin

Nephron segment	*n*	Tubular length per incubation (mm)	1α(OH)ase activity (fmol mm^{-1} h^{-1})
Glomerulus	4	50–100*	ND
Proximal convoluted tubules	4	21–38	0.65 ± 0.08
Proximal straight tubules	4	23–26	0.69 ± 0.09
Medullary thick ascending limb	4	39–85	ND
Cortical thick ascending limb	4	23–78	ND
Distal convoluted tubules	4	19–35	ND
Collecting tubules	4	32–62	ND

Defined nephron segments were dissected from kidney slices of vitamin D-deficient rats 8 h after four injections of calcitonin, 20 MRC U per kg every 2 h and assayed for 1α(OH)ase activities as described in Fig. 1 legend. Enzyme activities in kidneys of vitamin D-deficient rats without calcitonin treatment were 0.70 ± 0.05 fmol mm^{-1} h^{-1} in the PCT and undetectable in all other nephron segments including the PST[1]. Results are the mean ± s.e.m. ND, not detectable or < 1 fmol per incubation.

* The number of glomeruli per incubation.

---

* To whom correspondence should be addressed.

0028-0836/81/220327—03$01.00

Table 2 Effect of calcitonin on 1α(OH)ase in PCT and PST of thyroparathyroidectomized vitamin D-deficient rats

	1α(OH)ase activity (fmol mm^{-1} h^{-1}) PCT	PST
Control	0.17±0.10	ND
With calcitonin treatment	0.16±0.06	0.73±0.11

Rats were fed a vitamin D-deficient diet containing 0.45% calcium and 0.3% phosphorous for 4–6 weeks from weanling, then supplemented with additional 0.75% calcium (total 1.2% calcium) for 3–5 days before thyroparathyroidectomy done under light anaesthesia. Starting 16 h after surgery, the rats received four successive subcutaneous injections of either calcitonin, 50 MRC U per kg, and calcium gluconate, 200 mg per kg, or the vehicle (control) every 2 h; they were killed 24 h after the surgery. Results are the mean± s.e.m., *n* = 4–6. ND, not detectable or <0.05 fmol mm^{-1} h^{-1}.

tubular length or per glomerulus, were calculated from data obtained by TLC and the specific activity of ^3H-25(OH)D_3 corrected for recovery (82–92%) and counting efficiency (43–52%)[1]. The results were expressed as the mean ± s.e.m. and analysed by the *t*-test.

As shown in Fig. 1*a* and *b*, the administration of calcitonin to vitamin D-deficient rats resulted in a marked stimulation of 1α(OH)ase activity in PST which was undetected in control rats. This stimulatory effect of calcitonin appeared at 4–6 h and reached a plateau at 8 h (Fig. 1*a*); maximal stimulation was obtained with 20 MRC units per kg every 2 h (Fig. 1*b*). Table 1 shows the distribution of the enzyme activity along the nephron segments from hormone-treated rats. 1α(OH)ase activities in the PCT in rats with and without calcitonin treatment were 0.65±0.08 and 0.70±0.05 fmol mm^{-1} h^{-1}, respectively. There was no stimulation of the enzyme by calcitonin in other nephron segments tested, including glomeruli, medullary and cortical thick ascending limbs of Henle's loop, distal convoluted tubules and collecting tubules.

These results clearly demonstrate that calcitonin selectively stimulated 1α(OH)ase activity in the PST, and that the enzyme activity was undetectable in control vitamin D-deficient rats. However, these results do not exclude a possible stimulatory effect of calcitonin on the enzyme activity in the PCT, as it may have failed to stimulate 1α(OH)ase activity in the PCT because the activity was maximally activated by the hyperparathyroidism of vitamin D deficiency. To test this possibility, we examined the effect of calcitonin on the enzyme activity in both the PCT and PST of thyroparathyroidectomized vitamin D-deficient rats. As shown in Table 2, calcitonin stimulated the 1α(OH)ase in the PST but not in the PCT, indicating the specificity of the stimulation of 1α(OH)ase in the PST. Our results are in accord with *in vivo* observations by Horiuchi *et al.* that calcitonin stimulated the 1α(OH)ase independently of parathyroid hormone and that the effects of the two hormones are additive[4]. Furthermore, it is evident from our data that calcitonin and parathyroid hormone can act on anatomically and functionally distinct segments of the nephron.

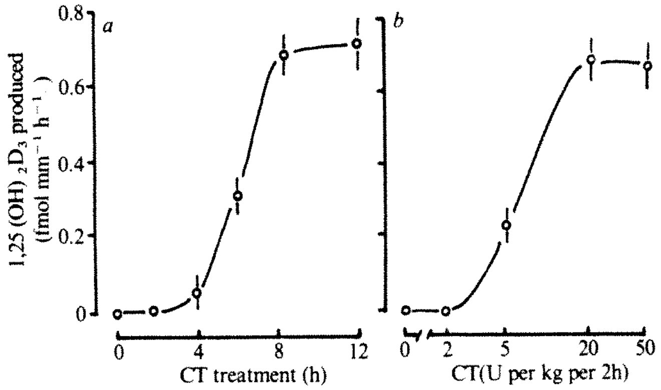


Fig. 1 Time course (*a*) and dose response (*b*) of the effects of calcitonin (CT) on the 1α(OH)ase activity in the PST of vitamin D-deficient rats. Under pentobarbital anaesthesia, 40 mg per kg, the left kidney was perfused with chilled Krebs–Ringer–bicarbonate buffer, *p*H 7.4, containing 8.3 mM glucose, 0.1% bovine serum albumin (fraction V, Sigma) and 0.1% collagenase (type I, Sigma). Kidney slices were then incubated for 30 min at 30°C in the same buffer with a constant bubbling with a gas mixture of 95:5% = O_2:CO_2. After rinsing the slices three times with 30–50 ml of ice-cold modified Hanks' solution (in mM: NaCl, 137; KCl, 5.2; Na_2HPO_4, 0.34; KH_2PO_4, 0.44; $CaCl_2$, 1.0; $MgCl_2$, 0.49; $MgSO_4$, 0.42; and $NaHCO_3$, 4.2; *p*H 7.4), the PST were dissected in the Hanks' solution at 0°C under a stereomicroscope. Tubular segments of known length were incubated for 60 min at 37°C in 20 µl of modified Hanks' solution containing 5 mM pyruvate, 8.3 mM glucose and 10 mM HEPES, *p*H 7.4, with 0.5 µM [26, 27-*Me* − ^3H]-25(OH)D_3 (specific activity 22.3 Ci $mmol^{-1}$) or [23, 24(n) − ^3H]-25(OH)D_3 (specific activity 102 Ci $mmol^{-1}$), both isotopes from Amersham) with room air as a gas phase as reported previously[1]. Lipid extracts[11] were applied to silica gel 60 TLC and ^3H − 1, 25(OH)$_2$D$_3$ produced, verified by HPLC[1], were calculated in fmol per mm tubular length per h based on the specific activity of ^3H − 25(OH)D_3, corrected for recovery (82–92%) and counting efficiency (43–52%). *a*, Calcitonin, 20 MRC U per kg, was given every 2 h. *b*, Calcitonin was given at doses indicated on the abscissa every 2 h and rats were killed at 8 h.

Parathyroid hormone is believed to stimulate 1α(OH)ase by stimulating adenylate cyclase[5,6]. As calcitonin also stimulates adenylate cyclase in the kidney, we examined the distribution of calcitonin-sensitive adenylate cyclase along the nephron both from vitamin D-deficient and normal, vitamin D-replete rats. As shown in Table 3, the hormone did not stimulate adenylate cyclase in the PST of either normal or vitamin D-deficient rats, but did stimulate the adenylate cyclase activity in the cortical thick ascending limb of Henle's loop, as has been reported previously[12]. These data strongly suggest that calcitonin may stimulate 1α(OH)ase in the PST of vitamin D-deficient rats by a mechanism independent of adenylate cyclase activation. The additive effects of CT and cAMP on 1α, 25(OH)$_2$D$_3$ production in thyroparathyroidectomized vitamin D-deficient rats[6] are consistent with such a postulate. Also, Larkins *et al.*[13] found that CT can stimulate 1α(OH)ase in chick kidney tubules *in vitro*.

Table 3 Calcitonin-sensitive adenylate cyclase activity along the rat nephron

	Adenylate cyclase activity, cyclic AMP produced (fmol per mm per 30 min) Normal		Vitamin D-deficient	
Nephron Segments	−CT	+CT	−CT	+CT
Proximal convoluted tubules	4.3±1.5	8.4±2.1	NT	NT
Proximal straight tubules	12.2±2.0	16.7±3.1	5.6±3.0	6.0±2.6
Medullary thick ascending limb	46.9±1.5	220.1±19.4*	NT	NT
Cortical thick ascending limb	40.9±19.3	433.6±54.6*	37.1±10.4	402.0±53.6*
Distal convoluted tubules	9.5±2.3	627.9±104.2*	NT	NT
Collecting tubules	11.5±0.6	75.0±4.7*	NT	NT

Adenylate cyclase activities in defined nephron segments were assayed with (+CT) or without (−CT) the addition of calcitonin in the incubation medium, by the method of Imbert *et al.*[14] with modifications[15]. Results are the mean ±s.e.m. of three to six incubations, and concentration of calcitonin used was 0.1 U ml^{-1} in all incubations except for PCT and PST of normal rats where it was 1.0 U ml^{-1}. NT, not tested.

*Compared with −CT, $P<0.01$.

Our results provide a plausible explanation for the presence of 1α(OH)ase in the PST of the kidney of the fetus[2] whose circulating calcitonin levels are probably elevated[7–10], but not in the kidney of mature, vitamin D-deficient rats[1], which are likely to have low plasma calcitonin levels. The data also suggest a possible physiological role, hitherto unrecognized, of calcitonin in vitamin D metabolism. We propose that calcitonin may be able to recruit additional tubular cells, that is those in the PST, to produce 1α, 25(OH)$_2$D$_3$; these cells are potentially capable of producing this sterol but may be latent in vitamin D deficiency and are recruited in special circumstances when the need for calcium is markedly enhanced, as may occur in the fetus.

We thank Dr Jack W. Coburn for valuable discussions. This work was supported by the Veterans Administration and grants from the USPHS (AM-21351 and AM-14750) and American

Diabetes Association. S. T. is an Advanced Research Fellow of American Heart Association, Greater Los Angeles Affiliate.

Received 16 December 1980; accepted 13 March 1981.

1. Kawashima, H., Torikai, S. & Kurokawa, K. *Proc. natn. Acad. Sci. U. S. A* **78,** 1199–1203 (1981).
2. Akiba, T., Endou, H., Koseki, C., Sakai, F., Horiuchi, N. & Suda, T. *Biochem. Biophys. Res. Commun.* **94,** 313–318 (1980).
3. Galante, L., Colston, K. W., MacAuley, S. J. & MacIntyre, I. *Nature* **238,** 271–273 (1972).
4. Lorenc, R., Tanaka, Y., DeLuca, H. F. & Jones, G. *Endocrinology* **100,** 468–472 (1977).
5. Horiuchi, N., Suda, T., Takahashi, H., Shimazawa, E. & Ogata, E. *Endocrinology* **101,** 969–974 (1977).
6. Horiuchi, N. *et al. Biochem. J.* **184,** 269–275 (1979).
7. Garel, J–M., Savajol, H. & Barlet, J.–P. *Biol. Neonate* **28,** 207–218 (1976).
8. Garel, J.–M, Martin-Rosset, W. & Barlet, J.–P. *Horm. Metab. Res.* **7,** 429–432 (1975).
9. Reynolds, W. A. *et al. Metabolism* **51,** 595–598 (1980).
10. Samaan, N. A., Anderson, G. D. & Adam-Mayne, M. E. *Am. J. Obstet. Gynecol.* **121,** 622–626 (1975).
11. Bligh, E. G. & Dyer, W. J. *Can. J. Biochem. Physiol.* **37,** 911–917 (1959).
12. Chabardes, D., Imbert-Teboul, M., Montegut, M., Clique, A. & Morel, F. *Proc. natn. Acad. Sci. U.S.A.* **73,** 3608–3612 (1976).
13. Larkins, R. G. *et al. Clin. Sci. molec. Med.* **46,** 569–582 (1974).
14. Imbert, M., Chabardes, D., Montegut, M., Clique, A. & Morel, F. *Pflügers Arch. ges. Physiol.* **354,** 213–228 (1975).
15. Torikai, S. & Imai, M. *Tohoku J. exp. Med.* **129,** 91–99 (1979).

Maintenance of blood pressure by the renin–angiotensin system in normal man

G. A. MacGregor, N. D. Markandu, J. E. Roulston & J. C. Jones

Department of Medicine, Charing Cross Hospital Medical School,
London W6 8RF, UK

J. J. Morton

MRC Blood Pressure Unit, Western Infirmary,
Glasgow G11 6NT, UK

The view that the renin–angiotensin–aldosterone system is a defence mechanism that maintains blood pressure only in conditions of volume depletion or sodium restriction[1] has been supported by studies with saralasin[2,3], a competitive inhibitor of angiotensin II, and teprotide[1], an inhibitor of angiotensin converting enzyme that blocks the formation of angiotensin II. However, both saralasin and teprotide will underestimate the role of the renin–angiotensin system because: (1) they have to be given intravenously, thus inhibiting only the short-term effects of circulating angiotensin II; (2) all subjects in these studies were rested for some time before infusion[1–3]; and (3) saralasin is an agonist[4–7] which only inhibits the action of circulating angiotensin II when its level is high. Niarchos *et al.*[8] demonstrated a drop in blood pressure with teprotide in normal subjects on a sodium intake of 150 mmol per day, but this was partly due to one subject fainting. Therefore, none of these studies clearly shows at what stage of volume depletion the system actively maintains blood pressure. By analogy with other biological systems, a gradual increase in the importance of the renin–angiotensin system as sodium is lost is more likely than an all-or-none phenomenon. Using the recently developed oral inhibitors of angiotensin converting enzyme[9–11], it is now possible to inhibit the formation of angiotensin II long-term and study human subjects during normal activity. We have now investigated the effect of one such inhibitor, captopril, in normotensive, healthy, male subjects on a normal sodium intake. Our results demonstrate that the renin–angiotensin system maintains blood pressure in these conditions and that the system is an important controller of aldosterone secretion as well as sodium excretion on a normal sodium intake.

A 3-week study was made of eight normotensive, healthy, male subjects (mean age 22 (21–26) yr), whose diastolic pressure was <80 mm Hg after a 2-month observation period. All subjects were on a fixed sodium intake of 120 mmol sodium per day and 115 mmol potasssium per day. During the 3 weeks of this diet all food and drink was supplied by the metabolic ward kitchen. The subjects were not admitted to hospital and were allowed to go about their normal activities although vigorous exercise was discouraged. On the eighth day of the diet, captopril therapy was started, and measurements were made 2 h after the first dose of 25 mg. They were then given 25 mg captopril, three times a day for the first day, increasing on the second day to three 50-mg doses, and on the third and fourth day to 100 mg three times a day. A final dose of 100 mg was given on the morning of the fifth day of captopril therapy. After stopping the captopril, the diet was continued for a further 9 days.

Table 1 Mean values for blood pressure, plasma angiotensin II, plasma renin activity and plasma aldosterone at various times during the study

	Before captopril	Within 2 h of first dose	Fifth day of captopril
Blood pressure (mm Hg)	85.3 ± 0.70	78.6 ± 2.2*	71.6 ± 1.9*
Plasma angiotensin II (pmol l^{-1})	25.3 ± 2.3	15.5 ± 1.9*	13.3 ± 1.7*
Plasma renin activity (ng ml^{-1} h^{-1})	3.3 ± 0.5	5.7 ± 0.7*	41.4 ± 7.1*
Plasma aldosterone (pmol l^{-1})	549 ± 64	184 ± 26*	158 ± 14*

Values are mean ± s.e.m.
*$P < 0.001$ by Student's *t*-test.

Daily measurements of blood pressure and 24-h urinary sodium and potassium excretion were made. Urinary aldosterone was measured 3 days before, during the 5 days of and for 3 days after captopril therapy. Blood pressure was measured using semi-automatic ultrasound sphygmomanometers with automatic recorders (Arteriosonde, Roche)[12] in the same room, by the same nurse, at the same time of day, except for the 2-h post-captopril measurement. The mean value of five readings at 1–2-min intervals was taken as the blood pressure in that position, except for post-exercise blood pressure, where a single reading was taken 1 min after a standard period of exercise on a treadmill. Pulse rate was measured using a Cambridge 3048 pulse monitor. Measurements of plasma concentrations of angiotensin II, aldosterone, electrolytes and creatinine, plasma renin activity and blood urea were made on the fifth and eighth day before captopril, 2 h after the first dose of captopril and on the fourth and fifth day during captopril treatment. Plasma noradrenaline was measured on the eighth day before, and the fifth day of captopril treatment. All blood was taken without stasis after sitting upright for 5–10 min. Plasma angiotensin II (ref. 13), plasma renin activity[14], plasma[15] and urinary alsosterone were measured by radioimmunoassay, and plasma noradrenaline by radioenzymatic assay[16]. All results are reported as mean ± s.e.m. Mean blood pressure was calculated as diastolic pressure + (1/3 × pulse pressure). Statistical analysis was performed using Student's *t*-test.

The results of blood pressure, and plasma angiotensin II, renin activity and aldosterone measurements are shown in

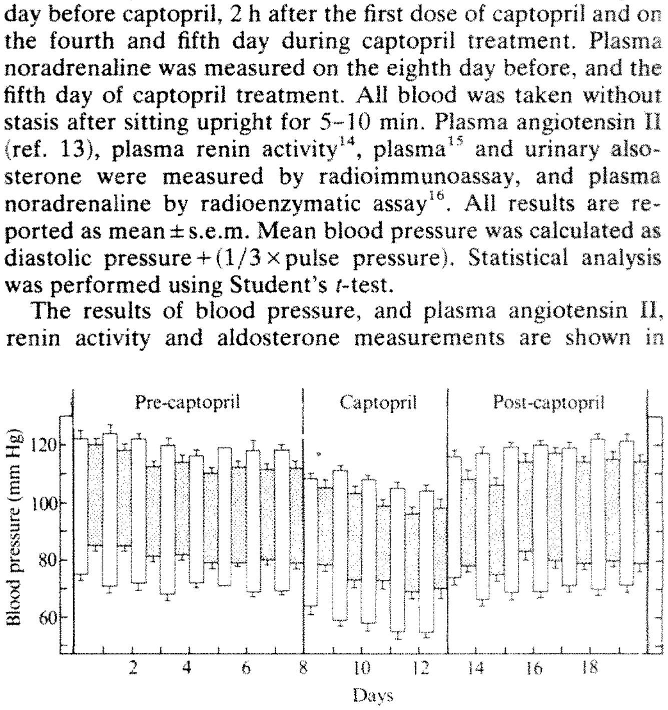


Fig. 1 Average daily supine (open bars) and standing (dotted bars) systolic and diastolic blood pressure (±s.e.m.) in the eight normal subjects on a sodium intake of 120 mmol per day, 7 days before, 5 days of and 7 days after captopril treatment. *2 h after first 25-mg dose of captopril.

0028–0836/81/220329—03$01.00

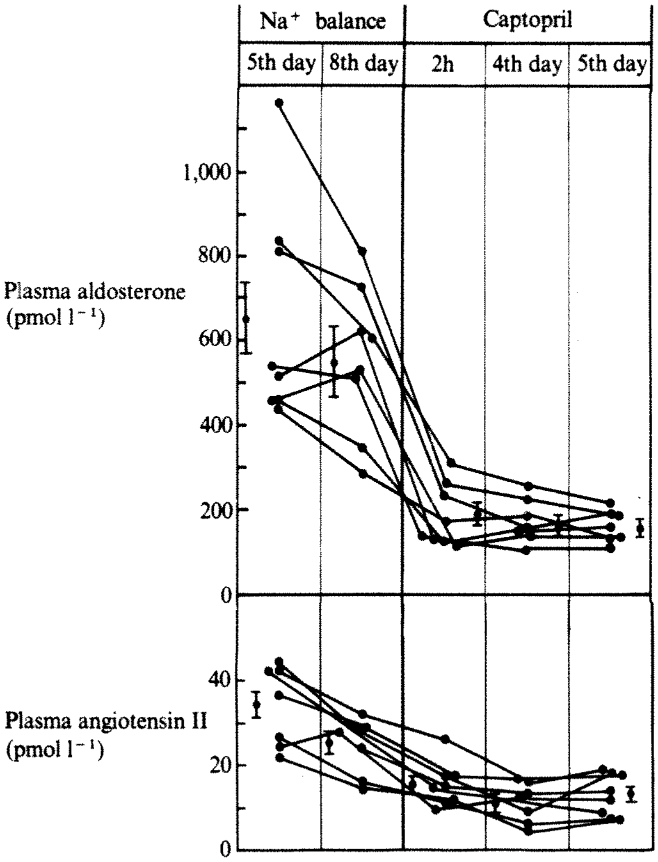


Fig. 2 Mean plasma aldosterone and angiotensin II in the eight normal subjects on a sodium intake of 120 mmol per day on the fifth and eighth day of sodium intake, 2 h after the first 25 mg of captopril and on the fourth and fifth day of captopril.

Table 1. After 5 days of captopril therapy, mean blood pressure had fallen by 16.5% compared with control values before captopril. On stopping captopril, blood pressure returned to the pre-captopril value. Figure 1 shows the average daily systolic and diastolic supine and standing pressure throughout the study. The fall in standing, sitting and post-exercise blood pressure was similar to that found with supine blood pressure. Pulse rate was not altered by captopril. Plasma renin activity rose, whereas plasma angiotensin II and aldosterone levels fell (see also Fig. 2).

Figure 3 shows the urinary aldosterone, potassium, sodium and weight for the 4 days before, the 5 days during and the 4 days after captopril therapy. There was a significant reduction in urinary aldosterone excretion during captopril therapy: mean values are 50.0(±4.7) nmol per 24h 3 days before captopril, and 36.5(±2.6) nmol per 24h ($P<0.01$) during captopril. All subjects were in a negative sodium balance during captopril treatment, with a mean loss of 142 mmol (±10.6) of sodium per subject. All subjects lost weight, by the fifth day of captopril, the mean weight loss was 1.15 kg(±0.31; $P<0.001$).

Plasma noradrenaline, which was 392(±39) pg ml^{-1} before captopril rose to 712(±125) pg ml^{-1} ($P<0.05$) by the fifth day of treatment. Blood urea, plasma electrolytes and plasma creatinine showed no significant change. Plasma potassium rose from 3.9(±0.1) to 4.1(±0.1) mmol l^{-1} by the fifth day of captopril, but this rise was not significant.

Thus, captopril given for 5 days caused a 16.5% fall in mean blood pressure in these normotensive, healthy subjects on a sodium intake of 120 mmol per day. There is controversy about the way in which captopril lowers blood pressure. The only known action of captopril is its inhibition of angiotensin converting enzyme—this was seen in our study by the fall in angiotensin II and aldosterone levels. Other studies have directly demonstrated this inhibition[10]. If, as seems likely, captopril lowers blood pressure through a fall in angiotensin II, our results indicate that the renin–angiotensin–aldosterone system is a normal mechanism for maintaining blood pressure in normotensive man. These findings with captopril are supported by the study of Niarchos *et al.*[8] using teprotide but contrast with those of Haber[1]. However both of these studies were short-term infusions in resting subjects.

Inhibition of angiotensin converting enzyme could lead not only to a fall in angiotensin II levels, but also to an accumulation of bradykinin, a potential vasodilator. Recent work has demonstrated that there is no increase in plasma bradykinin with captopril[17] or teprotide[18], although this does not necessarily exclude an increase in tissue bradykinin. Significant correlations between the fall in plasma angiotensin II[19,20] initial plasma renin activity[21,22] and the fall in blood pressure with captopril have been shown in patients with high blood pressure and in salt-depleted dogs[23]. However, the reduction in the direct vasoconstrictor effect of circulating angiotensin II may not be the only mechanism by which captopril lowers blood pressure[24]. Angiotensin II has many other actions; by acting on the area postrema in dogs, it can increase blood pressure[25]. Recent suggestions that the brain renin–angiotensin system may be an important mediator of the response to captopril may also be relevant[26]. The fall in plasma aldosterone that occurs with captopril could also contribute towards the fall in blood pressure both through the observed loss of sodium and water and by a postulated, but unproved, direct action of aldosterone on arteriolar tone.

Our results also demonstrate that the renin–angiotensin–aldosterone system is an important controller of plasma aldosterone on a normal sodium intake. Either the fall in plasma angiotensin II or aldosterone could account for the observed loss of sodium, so the system is also important in the maintenance of sodium balance on a normal sodium intake.

Finally, if angiotensin II in normotensive human subjects maintains blood pressure on a normal sodium intake, the renin–angiotensin system might be involved in the causation or maintenance of high blood pressure. It would seem likely from our results that, if the levels of angiotensin II in patients with high blood pressure are within the same range as those of normal subjects, the renin–angiotensin system maintains blood pressure to the same extent in hypertensives as in normotensive subjects[3]

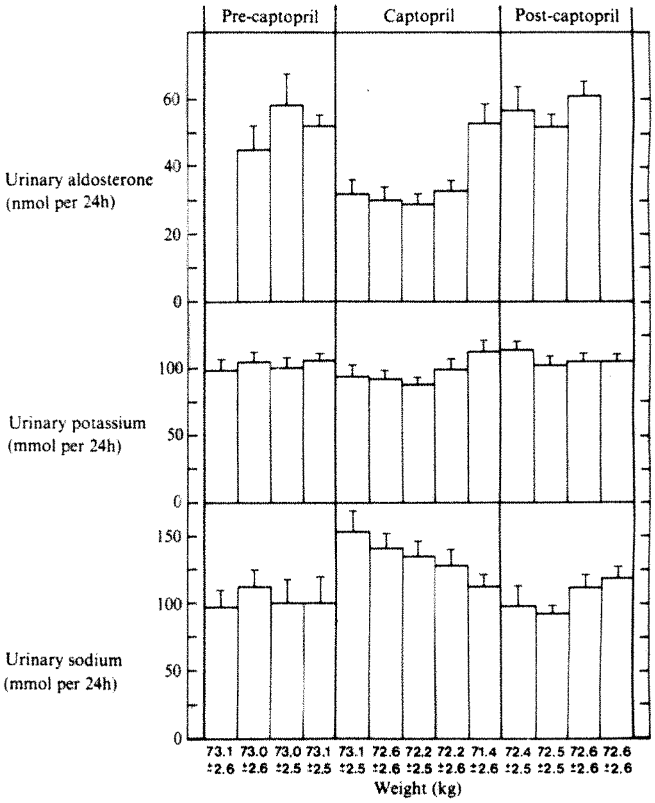


Fig. 3 Mean 24-h urinary aldosterone, potassium, sodium and weight (±s.e.m.) in the eight normal subjects for the 4 days before, 5 days of and 4 days after captopril treatment.

and may be an important mechanism whereby blood pressure can be lowered. However, the renin–angiotensin system can only be the cause of high blood pressure if the level of circulating angiotensin II is above normal, relative to the sodium balance, or there is an increased sensitivity to the actions of angiotensin II, compared with that in normotensive subjects.

We thank Professor P. Sever for measurements of plasma noradrenaline. This study was supported by the MRC and the Wellcome Trust.

Received 15 December 1980; accepted 5 March 1981.

1. Haber, E. *Circulation* **54,** 849–861 (1976).
2. Posternak, L., Brunner, H., Gavras, H. & Brunner, D. *Kidney International* **2,** 197–203 (1977).
3. MacGregor, G. A., Markandu, N. D. & Roulston, J. E. *Clin. Sci.* **57,** 145s–148s (1979).
4. MacGregor, G. A. & Dawes, P. M. *Br. J. clin. Pharmac.* **3,** 483–487 (1976).
5. Hollenberg, N. K., Williams, G. H., Burger, B. M., Ishikawa, I. & Adams, D. F. *J. clin. Invest.* **57,** 39–46 (1976).
6. Anderson, G. H., Streeten, D. H. P. & Dalakos, T. H. G. *Circulation* **40,** 243–250 (1977).
7. Case, D. B. *et al. Am. J. Med.* 790–796 (1976).
8. Niarchos, A. P., Pickering, T. G., Case, D. B., Sullivan, P. & Laragh, J. H. *Circulation Res.* **45,** 829–837 (1979).
9. Ondetti, M. A., Rubin, B. & Cushman, D. W. *Science* **196,** 441–444 (1977).
10. Ferguson, R. K., Turini, G. A., Brunner, H. R. & Gavras, H. *Lancet* **i,** 775–778 (1977).
11. Gavras, H. *et al. New Engl. J. Med.* **298,** 991–995 (1978).
12. George, C. F., Lewis, P. J. & Petrie, A. *Br. Heart J.* **37,** 804–807 (1975).
13. Dusterdieck, G. & McElwee, G. *Eur. J. clin. Invest.* **2,** 32–38 (1971).
14. Roulston, J. E. & MacGregor, G. A. *Clinica chim. Acta* **88,** 45–48 (1978).
15. James, V. H. T. & Wilson, G. A. *Methodological Developments in Biochemistry* Vol. 5 (ed. Reid, E.) 149–158 (Elsevier, Amsterdam, 1976).
16. Henry, D. P., Starman, B. J., Johnson, D. G. & Williams, R. H. *Life Sci.* **16,** 375–384 (1975).
17. Matthews, P. G., McGrath, B. P. & Johnson, C. I. *Clin. Sci.* **57,** 135s–138s (1979).
18. Hulthen, U. L. & Hokfelt, B. *Acta med. scand.* **204,** 497–502 (1978).
19. Fagard, P., Amery, A., Lijnen, P. & Reybrouck, T. *Clin. Sci.* **57,** 131s–134s (1979).
20. Atkinson, A. B. *et al. Clin. Sci.* **57,** 139s–143s (1979).
21. Laragh, J. H., Case, D. B., Atlas, S. A. & Sealey, J. E. *Hypertension* **2,** 586–593 (1980).
22. MacGregor, G. A., Markandu, N. D., Roulston, J. E. & Jones, J. C. *Br. med. J.* **2,** 1106–1109 (1979).
23. Morton, J. J., Tree, M. & Casals-Stenzel, J. *Clin. Sci.* **58,** 445–450 (1980).
24. Tree, M. & Morton, J. J. *Clin. Sci.* (in the press).
25. Joy, M. D. & Love, R. D. *Nature* **228,** 1303–1304 (1970).
26. Unger, T., Kaufmann-Buhler, I., Scholkens, B. & Ganten, D. *Eur. J. Pharmac.* (in the press).

Active immunization and passive transfer of resistance against sporozoite-induced malaria in infant mice

Augustine U. Orjih, Alan H. Cochrane & Ruth S. Nussenzweig

Division of Parasitology, Department of Microbiology,
New York University School of Medicine, 550 First Avenue,
New York 10016, USA

Immunization attempts using a variety of plasmodial preparations have identified several developmental stages and vaccination procedures which can protect against malaria[1–3]. These indicate that immunoprophylaxis against malaria may also be possible in man. The immune response to these vaccines has only been investigated in adult animals, but in countries where malaria is prevalent, children are most at risk of developing lethal infections. Thus we must ask whether vaccination against malaria would be effective when applied early in life. Here we report the use of intramuscular (i.m.) immunization of very young mice with radiation-attenuated sporozoites of *Plasmodium berghei* as a vaccination procedure, and that these animals can be effectively protected against sporozoite challenge. This approach avoids the intravenous (i.v.) immunization, which although highly effective in adult animals, is barely acceptable for use in man. Another important aspect of the host–parasite relationship in endemic areas is the role of immunity transmitted congenitally from mother to infants. We therefore investigated whether sporozoite-immunized adult female mice could transfer this immunity to their litters. It was found that the offspring acquired antisporozoite antibodies from their mothers through the milk. Furthermore, most of these offspring resisted challenge by infective sporozoites.

All experiments were performed in A/J mice using the NK65 strain of *P. berghei*. Infant mice were bred in our animal facility and the adults were obtained commercially (Jackson Laboratories, Bar Harbor). Sporozoites, collected from the salivary glands of laboratory-bred *Anopheles stephensi*, were γ-irradiated at 15,000 rad before being used in immunization[4].

Young mice, 2–24 days old, and adult mice, 8–10 weeks old, were inoculated i.m. with 7×10^4 irradiated sporozoites suspended in 0.1 ml of tissue culture medium (TC199). These animals received 1–3 additional immunizing doses of 3.5×10^4 irradiated sporozoites, each given by the same route and at intervals of 1–2 weeks. One to two weeks after the last immunizing dose, the immunized animals and age-matched, non-immunized controls were challenged with 1×10^4 infective sporozoites i.v., to test for protection. At the time of challenge, the young mice were 5–7 weeks old. Animals which failed to show parasitized erythrocytes in Giemsa-stained blood smears, 2 weeks after challenge, were considered to be protected.

The results of six experiments, involving the immunization and challenge of 74 young and 24 adult animals, are summarized in Table 1. A high percentage of the immunized young mice developed resistance to sporozoite challenge. The level of protection among the young animals varied from 71 to 100%. The highest per cent protection occurred in those animals which had received the first immunizing dose of sporozoites at the age of 7–14 days. Adult immunized mice had a significantly lower per cent protection. In each of the experiments, the control animals developed parasitaemia within 5 days after sporozoite challenge.

In addition to the demonstration that, in mice, sporozoite vaccination can be effectively administered early in life, our data suggest that the effectiveness of this vaccination varies with the age of the recipient animals. Adult mice are highly protected by i.v. immunization with sporozoites, but when the immunizing sporozoites are given i.m., the per cent protection against challenge decreases considerably[5]. In our experiments, only 33.3% of the adult animals were protected after i.m. immunization (Table 1).

Young animals did not produce detectable serum antibody levels in response to a single dose of irradiated sporozoites. However, after two and three immunizing doses, the young and adult mice had comparable levels of antibodies, as measured by the indirect immunofluorescent antibody test (Table 2).

Further experiments were done to determine whether sporozoite-immunized mice could transfer protection to their offspring; 70 mice corresponding to 13 litters born and nursed by sporozoite-immunized mothers, were challenged with infective sporozoites when the pups were between 18 and 25 days old (Table 3). The challenge had to be delayed until this age because a large proportion of mice <2 weeks old born to and nursed by non-immunized mothers, were found to be resistant to sporozoite-induced *P. berghei* infection (our unpublished

Table 1 Protection against *P. berghei* sporozoite-induced infection of very young and adult mice immunized i.m. with γ-irradiated sporozoites

Age of mice at the start of immunization	Challenged	Protected	Percentage of protection
Infants			
2–6 days	20	16	80.0
7–14 days	44	41	93.2
16–24 days	10	8	80
Adults	24	8	33.3
Non-immunized controls			
Infants	21	0	0
Adults	10	0	0

The data represent the results of six experiments. Irradiated sporozoites were administered 2–4 times, with intervals of 1–2 weeks, each animal receiving a total of 1.05–1.75×10^5 irradiated parasites during the immunization period. Non-immunized controls were age-matched with the immunized animals.

0028-0836/81/220331—02$01.00

Table 2 Serum antibody levels in young and adult animals after i.m. injection of γ-irradiated sporozoites

No. of immunizing doses of sporozoites	$\log_2$ anti-sporozoite antibody titre Pups	Adult mice
1	<4	5.0 ± 0.6
2	7.0 ± 0.5	7.0 ± 0.3
3	9.2 ± 0.2	9.0 ± 0.3

Antibody response as measured by indirect immunofluorescence of young and adult mice to either 1, 2 or 3 i.m. inoculations with irradiated *P. berghei* sporozoites. The primary immunizing dose consisted of 7×10^4 sporozoites; the second and third doses, given at weekly intervals, each consisted of 3.5×10^4 parasites. Serum was collected 2 weeks after completion of each immunization. Anti-sporozoite antibody titres are expressed as mean values ± s.e.m. for each group of 5 mice. The lowest titre considered as parasite-specific was 4. The experiment was repeated on three different occasions with similar results.

data). Such non-immunological resistance is possibly the result of an exclusive milk diet[6,7].

Of the pups born to and nursed by sporozoite-immunized mothers, 60% resisted the infective sporozoite challenge (Table 3). In contrast, age-matched mice born to and nursed by non-immunized mothers were almost uniformly (95%) susceptible to sporozoite challenge. The percentage of resistant offspring of immunized mothers varied from litter to litter. In one experiment, in which we simultaneously challenged four litters using the same sporozoite preparation, the range of protection varied from 28.5 to 100%.

Mice born to and nursed by sporozoite-immunized mothers showed minimal levels of antibodies at birth, which increased gradually with suckling time, reaching maximal levels at 2 weeks old. This progressive increase in antibody level with time suggests that mice acquire antibodies from their mothers largely through the milk. This was further supported by experiments in which mice born to non-immunized mothers were transferred to immune foster mothers. After 12 days of nursing by the latter, the pups had acquired an antibody level similar to that found in age-matched pups born to and nursed by sporozoite-immunized mice. Indirect immunofluorescence antibody titres, obtained at 13 and 21 days of age, were determined for individual serum samples and expressed as a mean value ± s.e.m. for each group of 5 mice. Titres of the immune mothers were usually a twofold higher dilution than those of the pups. Pups born to non-immunized mothers then transferred to immune foster mothers had $\log_2$ anti-sporozoite antibody titres of 8.6 ± 0.4 and 9.0 ± 0.5 at 13 and 21 days old, respectively, compared with titres of 9.0 ± 0.2 and 9.0 ± 0.3 for pups nursed by immunized mothers. The experiment was repeated three times, with similar results. Protection against sporozoite-induced infection in these animals also seems to be largely transferred through the milk; 55% of the animals born to non-immunized mothers and nursed by sporozoite-immunized mice resisted a sporozoite challenge, whereas all the offspring of immune mice, foster-nursed from their first day onwards by non-immunized mothers were susceptible to sporozoite infection (data not shown).

Table 3 Protection against *P. berghei* sporozoite-induced infection of pups born to and nursed by sporozoite-immunized mice

Mothers	Pups			
			Percentage of protection	
No.	No. challenged†	No. protected	Mean	Range
13 Immunized*	70	42	60	28.5–100
7 Non-immunized	22	1	4.5	0–20

The data represent the results of five experiments which included 13 litters of immunized mothers and 7 litters of non-immunized mothers.

*Mice were immunized with a total of 1.1×10^5 γ-irradiated sporozoites i.v., given in five doses before onset of pregnancy.

†Pups were 18–25 days of age when challenged i.m. with 2.1×10^4 infective sporozoites. This inoculum consistently resulted in patent infection in control animals.

The passively acquired antisporozoite antibodies largely belong to the IgG class. In contrast, mice actively immunized with sporozoites possess high levels of both IgG and IgM, as detected by immunofluorescence. Whereas the passive transfer of anti-sporozoite antibodies through the milk has been clearly documented in this system, our experiments do not rule out the concomitant transfer of immune cells, which could also contribute to the functional immunity detected in the offspring[8,9].

Regarding the role of antibodies in anti-sporozoite-mediated resistance to malaria infection, it has been recently demonstrated that monoclonal antibodies to *P. berghei* sporozoite surface antigen confer complete protection on recipient mice[10]. However, our experiments constitute the first instance in which the passive transfer of sporozoite-induced immunity was observed in natural conditions.

In view of these findings in rodents, it is certainly worth investigating whether passive transfer of sporozoite-induced immunity occurs in infant humans living in areas in which malaria is endemic.

We thank Ms Rita Altszuler and Ms Marilyn Maracic for technical assistance, and Ms Sharon Hecht-Ponger for typing the manuscript. A.U.O. is a recipient of a WHO training grant. This work was supported by the WHO and the Agency for International Development.

Received 5 January; accepted 12 March 1981.

1. Mitchell, G. H., Butcher, G. A. & Cohen, S. *Immunology* **29**, 397–407 (1975).
2. Gwadz, R. W. & Green, I. *J. exp. Med.* **148**, 1311–1323 (1978).
3. Nussenzweig, R. S. *Int. J. nucl. Med. Biol.* **7**, 89–96 (1980).
4. Orjih, A. U. & Nussenzweig, R. S. *Am. J. trop. Med. Hyg.* **29**, 343–347 (1980).
5. Spitalny, G. L. & Nussenzweig, R. S. *Proc. helminth. Soc. Wash.* **39**, 506–514 (1972).
6. Maegraith, B. G., Deegan, T. & Sherwood Jones, E. *Br. med. J.* **2**, 1382–1384 (1952).
7. Hawking, F. *Br. med. J.* **i**, 1201–1202 (1953).
8. Beer, A. E., Billingham, R. E. & Head, J. *J. invest. Derm.* **63**, 65–74 (1974).
9. Trentin, J. J., Gallagher, M. T. & Priest, E. L. *Transplantn. Proc.* **9**, 1473–1475 (1977).
10. Potocnjak, P., Yoshida, N., Nussenzweig, R. & Nussenzweig, V. *J. exp. Med.* **151**, 1504–1513 (1980).

Inducer T lymphocytes synthesize a factor that stimulates proliferation of cloned mast cells

Gary Nabel, Stephen J. Galli, Ann M. Dvorak, Harold F. Dvorak & Harvey Cantor

Departments of Cell Biology and Pathology, Sidney Farber Cancer
Institute, Departments of Pathology, Beth Israel Hospital and
Harvard Medical School, and the Charles A. Dana Research
Institute, Beth Israel Hospital, Boston, Massachussetts 02215, USA

Inducer T lymphocytes activate other cells to divide and express new function. Known target cells include other lymphocytes and haematopoietic stem cells[1]. We now provide evidence that the inducer T cell acts on another important target population: mast cells. Mast cells have a central role in the expression of immediate hypersensitivity and are also prominent in T-cell mediated reactions of the delayed type[2–7]. Because the proliferation of differentiated cells is often regulated by soluble growth factors, we examined an inducer T-cell clone for its ability to stimulate mast cell proliferation. We report here that cloned $Ly1^+2^-$ inducer T cells produce a factor that selectively induces morphologically and karyotypically normal mouse mast cell clones to proliferate. We therefore suggest that inducer T cells may regulate mast cell numbers by releasing a soluble growth factor that stimulates them to divide. Because mast cell products also affect certain T-cell functions[8–10], mast cell–T cell interactions may comprise part of an immunoregulatory circuit.

Mast cells proliferate in the thymus of NZB mice[11] and elsewhere in association with T cells[2–7]. To explain these findings, Burnet proposed that certain mast cell populations are

Table 1 Newly synthesized histamine content of mast cell or T-cell clones

Clone	Morphology	Newly synthesized histamine (pg per 10^6 cells per 19 h)
Cl. MC/8	Mast cell	4.5×10^4
Cl. MC/9	Mast cell	2.8×10^4
Cl. TL$^+$Ly1$^+$2$^+$/11	Lymphocyte	0
Cl. Ly1$^-$2$^-$/11	Lymphocyte	0

Histamine synthesis of two mast cell (Cl. MC/8 and 9) and two T-cell clones was measured by isotopic TLC[16].

derived from cells of the thymus-dependent lineage[12]. Alternatively, proliferation might be controlled by a soluble growth factor elaborated by T cells. We evaluated this question further using homogeneous populations of mast cells. Cells derived from the liver of a 13-day old mouse fetus from an A/J female mated with a C57BL/6 male were incubated in media conditioned by concanavalin A (Con A)-activated BALB/c spleen cells[1], and 10 days later were distributed at limiting dilutions in wells containing irradiated (2,000R) syngeneic bone marrow cells. Colonies appeared after 10–14 days with a cloning efficiency of ~10%. No growth of mast cells was observed from preparations of irradiated bone marrow cells. All colonies grown in these conditions were composed of cells of similar appearance that resembled mast cells by light microscopy because of their prominent metachromatic granules and expressed identical profiles of surface membrane glycoproteins (see below). These lines were subsequently cloned by micromanipulation[13], and have been designated Cl. MC/1-14. The cloned cells have normal karyotypes and were cultivated in large numbers ($>10^8$) with doubling times of 36–48 h. Clones can be maintained stably in the absence of any other cell type for at least 5 months, although doubling times may increase. Cl. MC/1-14 cells were further identified as mast cells by electron microscopy (Fig. 1) and by their ability to synthesize and store histamine (Table 1); histamine content ranged from 0.2 to 1.4 pg per cell. Preliminary studies show that cultured mast cells degranulate and release histamine when exposed to non-cytolytic concentrations (1 μg ml^{-1}) of the calcium ionophore A23187 or after passive sensitization with IgE and antigenic challenge. In addition, cloned cells can be frozen for storage and thawed with complete recovery of function and phenotype.

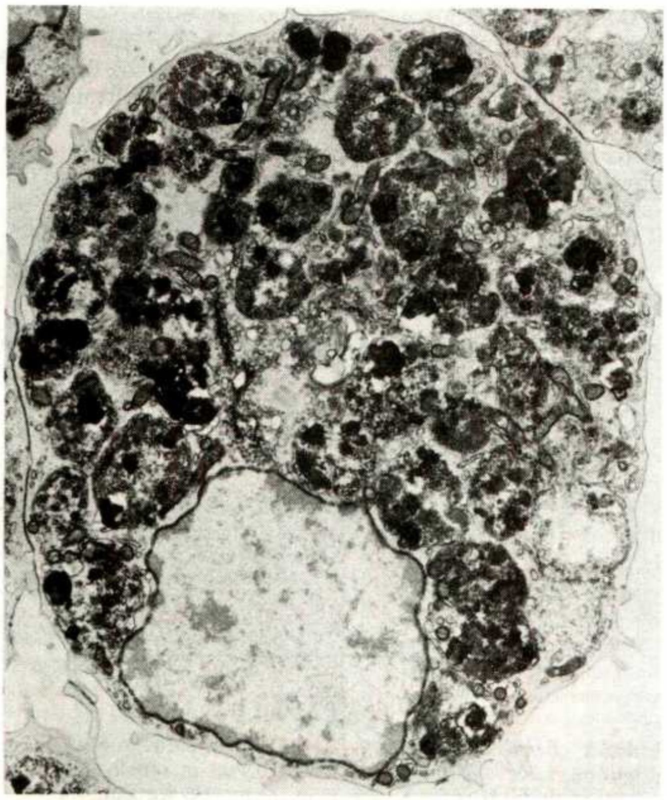


Fig. 1 Electron micrograph of a typical mast cell from Cl. MC/8. Numerous characteristic cytoplasmic granules contain varying amounts of electron-dense particles and small vesicles. By ultrastructural criteria, cloned mast cells closely resemble the mast cells present in murine connective tissue, lymphoid tissue and serous spaces[5]. Cells were fixed and processed by the osmium potassium ferrocyanide method[17]. Lead citrate stain, ×7,000.

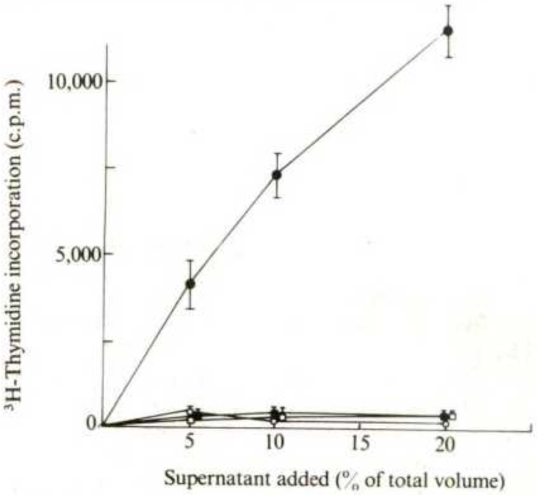


Fig. 2 Proliferation of cloned mast cells in the presence of supernatants of cloned inducer T cells. Cl. MC/9 cells, final concentration of 10^5 cells ml^{-1}, were incubated in flat-bottom microplate wells (Falcon 3040) in 0.1 ml of modified Dulbecco's modified Eagle's medium[1] with 4% heat-inactivated fetal calf serum, 5×10^{-5} M 2-mercaptoethanol and 2 mM glutamine. These cultures were incubated with increasing concentrations of supernatants from different T-cell clones: Cl. Ly1$^+$2$^-$/9 (●), Cl. Ly1$^+$2$^+$/4 (○), Cl. Ly1$^-$2$^+$/16 (■) or Cl. Ly1$^-$2$^-$/11 (□). After incubation for 24 h at 37 °C in 10% CO_2, 0.5 μCi of ^{3}H-thymidine, 5.4 Ci mM^{-1} (NEN), was added. Four hours later, cultures were collected on to glass filter strips and counted by liquid scintillation. Supernatants were prepared as follows: cloned T cells were washed three times and resuspended at a final concentration of 5×10^5 cells ml^{-1} in modified Dulbecco's modified Eagle's media supplemented with 5 μg ml^{-1} human transferrin, 5 μg ml^{-1} bovine insulin, 5×10^{-5} M 2-mercaptoethanol and glutamine (2 mM). After incubation for 24 h at 37 °C in 10% CO_2, supernatant was collected (300g for 10 min) and filtered (0.22 μm Swinnex).

Immunofluorescence studies[1] indicate that Cl. MC/1-14 cells do not carry the Thyl, Ly1 and Ly2 surface membrane glycoproteins characteristic of T lymphocytes. However, they do express the Ly5 glycoprotein, as do lymphocytes and haematopoietic cells[14]. Mast cells therefore display a different profile of cell-surface glycoproteins from thymocytes or T lymphocytes. We cannot exclude the possibility that thymic lymphocytes and mast cells arise from a common precursor because prothymocytes, like the mast cells described here, express the Thy1$^-$Ly5$^+$ glycoprotein profile[15]. Finally, the ultrastructure and low histamine content of some clones suggest that such clones may represent immature mast cells.

Mast cell clones divide continuously in media conditioned by Con A-activated spleen cells, but not in non-conditioned media. The medium of spleen cells incubated in the absence of Con A does not support growth. To define the cellular source of the mast cell growth factor, conditioned medium was prepared from spleen-derived T-lymphocyte clones[1]. Ly1$^+$2$^-$ inducer T-cell clones, in contrast to other T-cell clones, released a factor into the culture medium which stimulates mast cell proliferation (Fig. 2). The release of this factor is enhanced further by addition of Con A. This material is retained after dialysis, inactivated by pronase and has an apparent molecular weight of 40,000–45,000 on Sephacryl S-200. The partially purified factor does not stimulate significant proliferation of cloned T cells or fibroblasts (data not shown).

We have shown that pure mast cell populations may be readily and selectively derived from mouse haematopoietic tissue under the influence of an inducer T-cell product. This work is of interest for several reasons. First, such clones should prove useful for investigations requiring purified mast cells, particularly because alternative sources (such as peritoneal cavity mast cells) comprise a heterogeneous population that provides small cell numbers. They may also serve to test whether all mast cells require T-cell signals to the same degree[7]. Second, cloned lines allow definitive characterization of effector and target cell populations. For example, we have identified the mast cell as a direct cellular target of inducer T cells. This effect is not mediated through small numbers of an intervening cell, such as the macrophage, often present in highly enriched T-cell populations. Third, the data provide direct evidence to support earlier observations *in vivo* that link lymphocytes to mast cell hyperplasia in the thymus and at sites of delayed hypersensitivity. It will also be of interest to determine whether inducer T cells influence the growth of mastocytomas because these mast cell tumours commonly include large numbers of lymphocytes[6,7].

Although we have shown that T lymphocytes induce mast cell proliferation, others have demonstrated that mast cell products depress certain T lymphocyte functions. Thus, histamine may inhibit *in vitro* generation of cytotoxic T lymphocytes[8], macrophage inhibitory factor production[9], and cytotoxic T-lymphocyte effector function[10]. Taken together, these findings point to mutual interactive relationships between T lymphocytes and mast cells. These cellular interactions may explain mast cell hyperplasia in a variety of clinical and experimental hypersensitivity settings and suggest a control mechanism to curtail such reactions.

We thank Justine Osage and Rachel Pocinki for technical assistance. This work was supported by USPHS grants GM 07753, AI 12184, AI 13600, CA 28834 and AI 16925.

Note added in proof: Cl.Ly1^+2^- mass cell growth factor has been purified to apparent homogeneity. It has a p*I* of ~6.0 and $_r$ of 45,000 (refs 1, 18).

Received 27 October 1980; accepted 10 March 1981.

1. Nabel, G., Greenberger, J. S., Sakakeeny, M. A. & Cantor, H. *Proc. natn. Acad. Sci. U.S.A.* **78,** 1157–1161 (1981).
2. Dvorak, H. F. *et al. Lab. Invest.* **31,** 111–130 (1974).
3. Dvorak, A. M., Mihm, M. C. Jr & Dvorak, H. F. *Lab. Invest.* **34,** 179–191 (1976).
4. Ishizaka, T., Okudaira, H., Mauser, L. E. & Ishizaka, K. *J. Immun.* **116,** 747–754 (1976).
5. Ginsburg, H., Nir, I., Hammel, I., Eren, R., Weissman, B. A. & Naot, Y. *Immunology* **35,** 485–502 (1978).
6. Galli, S. J. & Dvorak, H. F. in *Cellular, Molecular, and Clinical Aspects of Allergic Disorders* (eds Gupta, S. & Good, R. A.) 1–53 (Plenum, New York, 1979).
7. Askenase, P. W. *Springer Semin. Immunopath.* **2,** 417–442 (1980).
8. Schwartz, E., Sutton, S., Askenase, P. W. & Gershon, R. K. *Immunopharmacology* **2,** 179 (1980).
9. Rocklin, R. E. *J. clin. Invest.* **57,** 1051–1058 (1976).
10. Plaut, M., Lichtenstein, L. M., Gillespie, E. & Henney, C. S. *J. Immun.* **111,** 389–394 (1973).
11. Burnet, F. M. *J. Path. Bact.* **89,** 271–284 (1965).
12. Burnet, F. M. *Cell. Immun.* **30,** 358–360 (1977).
13. Nabel, G., Fresno, M., Chessman, A. & Cantor, H. *Cell* **23,** 19–28 (1981).
14. Scheid, M. P. & Triglia, D. *Immunogenetics* **9,** 423–433 (1979).
15. Silverstone, A. E. *et al.* in *Differentiation of Normal and Neoplastic Hematopoietic Cells* (eds Clarkson, B., Marks, P. & Till, J.) 433–453 (Cold Spring Harbor Laboratory, New York, 1978).
16. Galli, S. J., Galli, A. S., Dvorak, A. M. & Dvorak, H. F. *J. Immun.* **117,** 1085–1092 (1976).
17. Dvorak, A. M., Hammond, M. E., Dvorak, H. F. & Karnovsky, M. J. *Lab. Invest.* **27,** 561–574 (1972).
18. Nabel, G. *et al.* (in preparation).

In vivo effect of anti-asialo GM1 antibody on natural killer activity

Masataka Kasai*, Takahiro Yoneda*,
Sonoko Habu†, Yoneo Maruyama*,
Ko Okumura‡ & Tohru Tokunaga*

* Department of Tuberculosis, National Institute of Health,
Shinagawa-ku, Tokyo 141, Japan
† Department of Pathology, Tohkai University School of Medicine,
Isehara, Kanagawa 259-11, Japan
‡ Department of Immunology, Faculty of Medicine, University of
Tokyo, Bunkyo-ku, Tokyo 113, Japan

Mounting evidence suggests that natural killer cells (NK) play an important part in the destruction of incipient tumours. If that is so, the destruction of NK cells *in vivo* should allow tumour cells to proliferate. We have recently described an antiserum having anti-NK activity which was raised against mouse brain tissue, and its effective antibody was shown to be directed against the cell surface glycolipid GM1 with its sialic acid groups removed (asialo GM1)[1]. We show here that intravenous (i.v.) injection of microlitre amounts of anti-asialo GM1 into BALB/c *nu/nu* mice almost completely abolishes NK activity and enhances the local growth of syngeneic lymphoma, RL♂1 as well as the incidence of tumour take. Evidence suggests that this effect is due to membrane damage producing loss of NK function rather than to cytolytic membrane attack.

The anti-asialo GM1 used in this study was prepared as an ammonium sulphate precipitate of whole antiserum of rabbits immunized with asialo GM1 as described previously[1]. As shown in Fig. 1, *in vivo* administration of anti-asialo GM1 to BALB/c nude mice resulted in a marked reduction in NK activity. The amount of anti-asialo GM1 required for elimination of NK activity was much smaller than expected. The dose response curve clearly indicates that even a few microlitres of anti-asialo GM1 is enough to eliminate almost all the NK activity in nude mouse spleen cells. Similar results were also obtained in various lymphoid organs (data not shown). The repressed state of NK activity lasted for several days, up to 1 week after one injection of the antiserum. Thus, we thought that anti-asialo GM1 could be used to study the biological role of NK cells *in vivo*. Table 1 shows that i.v. injection of anti-asialo GM1 into BALB/c nude mice greatly enhanced the local growth of syngeneic lymphoma, RL♂1 as well as the incidence of tumour take. This strong association between NK function and resistance to tumour growth agrees with the previous result which indicates the retardation of tumour growth by positively selected NK cells[2]. There are also remarkable parallels between the present results

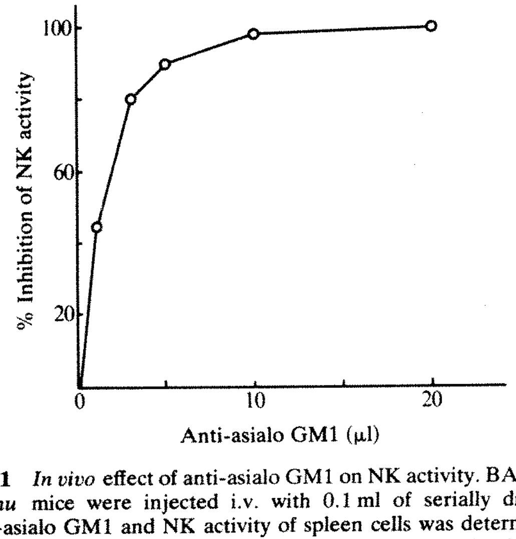


Fig. 1 *In vivo* effect of anti-asialo GM1 on NK activity. BALB/c *nu/nu* mice were injected i.v. with 0.1 ml of serially diluted anti-asialo GM1 and NK activity of spleen cells was determined after 24 h. The immunization protocol and determination of specificity for anti-asialo GM1 antiserum are described elsewhere[1]. NK activity was determined as follows: 2×10^4 ^{51}Cr-labelled YAC-1 cells (a Molony virus-induced lymphoma in A/Sn mice) were incubated in RPMI 1640 medium plus 10% fetal calf serum with (1×10^6) spleen cells at 37 °C in 5% CO_2. The amount of γ radioactivity released from triplicate cultures after 4 h was determined.

0028-0836/81/220334—02$01.00

Table 1 Effect of anti-asialo GM1 on tumour takes in BALB/c nude mice

Injected with	Frequency of tumour takes* Days after inoculation: 0	4	8	16	20	24
NRS	0/20	0/20	3/20	5/20	5/20	6/20
Anti-asialo GM1	0/20	16/20	16/20	20/20	20/20	20/20

BALB/c *nu/nu* mice were injected i.v. with 0.1 ml of anti-asialo GM1 (diluted 1:10) and then 2×10^6 RL♂1 cells were inoculated subcutaneously. The same amount of anti-asialo GM1 was injected repeatedly at intervals of 3 days. As a control, the same amount of normal rabbit serum (NRS) instead of anti-asialo GM1 was injected.

* No. of animals showing tumour growth out of the total no. grafted.

and the recent observation indicating that the growth of transplanted tumour cells is enhanced in beige mice with impairment of NK activity[3,4]. Furthermore, subsequent studies have shown that anti-asialo GM1 could be of practical value in improving the transplantation of human tumour cells which do not readily take even in nude mice[5].

As the reduction of NK activity was also observed even 1 h after injection of anti-asialo GM1, we set up an experiment to determine the mode of action of anti-asialo GM1 on NK activity. As can be seen in Table 2, the incubation of BALB/c *nu/nu* spleen cells with anti-asialo GM1 for 1 h at 37°C resulted in a strong reduction in NK activity only when normal mouse serum (NMS) was present. As was the case *in vivo*, NK activity was inhibited by even a very small amount of anti-asialo GM1, whereas NMS alone was not effective. Because of the weak activity of mouse complement, it seems unlikely that abolition of NK activity is due to the cytolytic effect caused by microlitre amounts of anti-asialo GM1. In fact, reduction of NK activity was caused by serum from AKR mice with a genetically determined deficiency of C5 as well as by serum from mouse strains possessing the complete spectrum of complement components (Table 2). Serum from C4-deficient guinea pigs was also effective in removing NK activity in the presence of anti-asialo GM1. These results indicate that the cytolytic membrane attack pathway is not required for this reaction. It seems reasonable to assume that anti-asialo GM1 antibody and heat-labile factor(s) in NMS cause a certain membrane damage which gives rise to loss of NK function, rather than to the cytolytic membrane attack pathway.

In any event, the distinct *in vivo* effect of anti-asialo GM1 in nude mice can probably help us to understand some aspects of natural defense mechanism against neoplastic growth. Furthermore, it is quite conceivable that a very small amount of autoantibody to asialo GM1 could be the key immunoregulatory negative feedback signal for NK activity.

Table 2 Requirement of serum factor for elimination of NK activity by anti-asialo GM1 *in vitro*

Experiment	Treatment of BALB/c *nu/nu* spleen cells	NK activity (% lysis)
a	Anti-asialo GM1 alone	32.1 ± 1.5
	NMS (BALB/c *nu/nu*) alone	31.4 ± 1.1
	NMS + anti-asialo GM1	0.5 ± 0.3
	NMS (56 °C, 30 min) + anti-asialo GM1	30.7 ± 1.2
b	NMS (AKR) alone	31.2 ± 1.5
	NMS + anti-asialo GM1	0.8 ± 0.4
c	GPC (C4-deficient guinea pig serum) alone	30.0 ± 0.9
	GPC + anti-asialo GM1	1.1 ± 0.8

NK activity of BALB/c *nu/nu* spleen cells was determined after the following treatments. *a*, Spleen cells (5×10^6 per 0.1 ml) were incubated for 1 h at 37 °C with anti-asialo GM1 (1:100) diluted in the serum of BALB/c *nu/nu* mice. *b*, Spleen cells were incubated for 1 h at 37 °C with anti-asialo GM1 (1:100) diluted in the serum of AKR mice. *c*, Spleen cells were incubated for 1 h at 37 °C with anti-asialo GM1 (1:100) diluted in C4-deficient guinea pig complement (GPC). Effector/target ratio was 50:1.

We thank Dr H. Okada for C4-deficient guinea pig serum and Ms N. Katsu for secretarial help.

Received 11 November 1980; accepted 9 March 1981.

1. Kasai, M., Iwamori, M., Nagai, Y., Okumura, K. & Tada, T. *Eur. J. Immun.* **10,** 175–180 (1980).
2. Kasai, M., Leclerc, J. C., McVay-Boudreau, L., Shen, F. W. & Cantor, H. *J. exp. Med.* **149,** 1260–1264 (1979).
3. Kärre, K., Klein, G. O., Kiessling, R., Klein, G. & Roder, J. C. *Nature* **284,** 624–626 (1980).
4. Talmadge, J. E., Meyers, K. M., Prieur, D. J. & Starkey, J. R. *Nature* **284,** 622–624 (1980).
5. Habu, S. *et al.* (in preparation).

Interleukin-2 augments natural killer cell activity

Christopher S. Henney, Kagemasa Kuribayashi, Donald E. Kern & Steven Gillis

Basic Immunology Program, Fred Hutchinson Cancer Research Center, Seattle, Washington 98104, USA

The lymphokine interleukin-2 (IL-2; also termed T-cell growth factor) causes the proliferation of activated T-cell clones[1,2]. Effector T-cell lines exhibiting suppressor, helper and cytotoxic activities have been established by growth of appropriately stimulated cells in IL-2-containing medium[3,4]. In addition, IL-2 has been shown to provide requisite T-cell 'help' in a number of *in vitro* immune response assays, including the generation of cytotoxic T cells from thymocyte (and nude spleen cell) cultures[5] and the induction of erythrocyte-specific antibody in T cell-deficient populations[6]. We report here a further activity of IL-2, one previously attributed solely to interferon—the ability to augment the cytotoxic activity of natural killer (NK) cells. This conclusion was reached from the observation that preparations of IL-2, which lacked interferon, caused a rapid increase in NK cell activity. This ability to augment cytotoxity was removed from IL-2 preparations, not only by absorption with IL-2 receptor bearing cells, but also by precipitation with a monoclonal antibody directed against IL-2.

Our attention was originally drawn to IL-2 as an NK cell potentiator when we observed that spleen cells from adult C57BL/6, CBA/J and BALB/c nude (nu^+/nu^+) mice showed markedly increased NK cell activity when cultured for short (up to 24 h) periods in medium containing IL-2. As previously reported[7], murine NK cell reactivity was relatively labile in conventional culture conditions, and decreased ~60–80% after culture for 24 h. In contrast to this decline, cells cultured in medium containing IL-2 showed markedly increased NK cell reactivity, even when compared to the activity of fresh uncultured spleen cell populations. Increases in splenocyte NK cell reactivity were first observed after ~16 h of culture in IL-2-containing medium. The level of augmentation observed peaked after 24–30 h of culture and declined to startling levels by ~96 h.

Analysis of a variety of IL-2 preparations for conventional IL-2 activity (ability to support the proliferation of an IL-2-dependent cell line, as measured in a 24-h microassay[8] and concurrently for NK augmenting activity (ability to augment NK cell activity during a 24-h culture period; and subsequently assessed in a 4-h ^{51}Cr release assay[9]), showed a positive correlation between the two activities (Table 1).

It has been widely observed that interferon augments NK-cell reactivity[10,11], but several observations led us to believe that the augmentation caused by the IL-2 preparations was not due to contaminating interferon. First, in those experiments in which IL-2 preparations were added to graded doses of a purified mouse interferon preparation, augmentation of NK-cell reactivity was observed over and above that elicited by interferon alone (Fig. 1). In the conditions used, 10^3 units ml^{-1} interferon had maximal stimulatory effects, with amounts in excess of this

0028–0836/81/220335—04$01.00

Table 1 Effect of IL-2 containing preparations on murine NK-cell activity

Effector cells	Culture conditions	IL-2 ($U\ ml^{-1}$)	% Specific cytolysis against: YAC-1	L5178Y cl 27v	L5178Y cl 27av
BALB/c (nu^+/nu^+) spleen	Not cultured	—	21.7	14.7	1.1
	Medium alone	0	6.5	1.9	0.9
	Rat IL-2[a]	2.2	60.0	40.6	5.8
	PAGE-purified mouse IL-2	3.5	72.0	ND	ND
	Medium alone	0	3.3	0.1	0.8
	Rat IL-2[b]	0.6	24.2	12.7	1.3
	Syngeneic MLC supernatant ($C57BL/6+C57BL/6_{Mit}$)	0.05	9.6	5.2	2.9
	Congeneic MLC supernatant ($B10.BR+AKR/J_{Mit}$)	1.5	38.9	20.4	5.1
	Allogeneic MLC supernatant ($C57BL/6+CBA/J_{Mit}$)	2.4	50.5	27.7	7.6
	Medium alone	0	1.5	ND	ND
	Isoelectrically focused IL-2:				
	pI 4.3	1.7	23.6	ND	ND
	pI 4.9	1.5	22.7	ND	ND

10^7 spleen cells were cultured for 24 h at 37 °C in 2 ml of RPMI 1640 medium containing 10% fetal calf serum and 5×10^{-5} M 2-mercaptoethanol. At the end of culture, the cytotoxic activity was assessed in a 4-h ^{51}Cr-release assay with YAC-1, cl 27v and cl 27av as target cells. A range of effector : target cell ratios was used; the data shown are for multiplicities of 50 : 1. Cl 27v and cl 27av are both cell lines established from the DBA/2 lymphoma L5178Y which differ in their susceptibility to NK cells, cl 27v being susceptible, cl 27av being insusceptible[20]. Both of these lines are lysed to an equivalent extent by alloimmune cytotoxic T cells[20]. The cytotoxic activity measured above was attributed to NK cells by a number of criteria, but notably by the ability to lyse 27v cells but not 27av, and by the observation that the lytic activity was abrogated by pretreatment of effector cells with complement and an antiserum directed against the neutral glycolipid, asialo GM_1 which is displayed on NK cells but not on cytotoxic T lymphocytes[21]. The IL-2 preparations used were: rat IL-2 supernatants from 24 h Con A ($5\ \mu g\ ml^{-1}$) stimulated rat spleen cells (10^7 cells ml^{-1}) used either at a dilution of 1 : 20 or diluted 1 : 80; purified mouse IL-2 supernatant from PHA-stimulated LBRM lymphoma cells[14] eluted off a gel slice from a polyacrylamide gel (PAGE) or from an isoelectric focusing gel, as indicated (for experimental details of preparations see refs 13, 14); murine MLC supernatants (diluted 1 : 2) from 24-h culture of 10^7 responder spleen cells and 10^7 mitomycin-C treated congenic or allogenic splenocytes as indicated. ND, not determined. Similar findings to those shown were also made using CBA/J spleen cells as a source of NK cells.

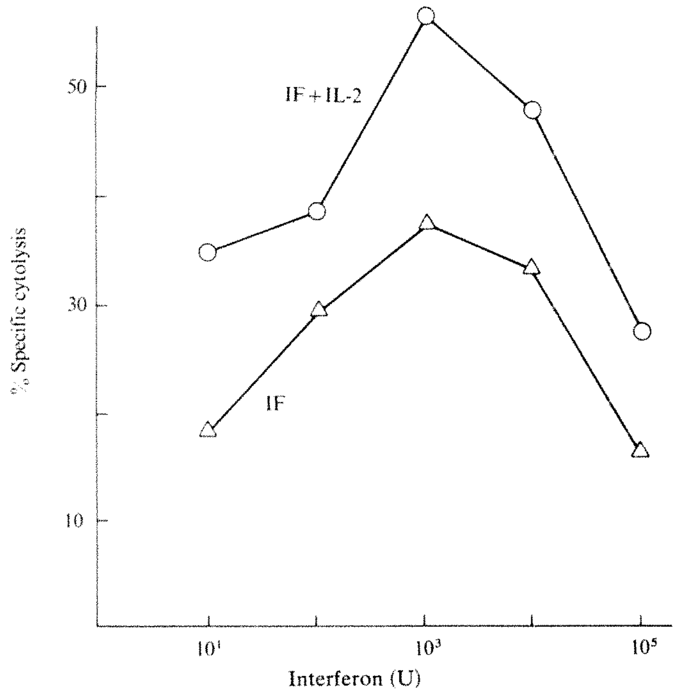


Fig. 1 Effects of interferon (IF) and interleukin-2 (IL-2) on the activity of CBA/J splenic NK cells. Various amounts of purified mouse fibroblast IF (1.1×10^7 units per mg protein, Calbiochem-Behring) were added to 10^7 normal CBA/J spleen cells and incubated for 24 h at 37 °C in the presence or absence of 1 unit of purified murine IL-2 (prepared by sequential $(NH_4)_2SO_4$ precipitation, ion-exchange chromatography and flat bed isoelectric focusing of mitogen-induced mouse spleen cell conditioned medium using methods described elsewhere[16]. After 24 h, the NK-cell mediated lytic activity of the cultures was assessed by a 4-h ^{51}Cr release assay using YAC-1, L5178Y cl 27v and 27av target cells, at various effector: target cell ratios. The data shown are for YAC-1 target cells and an effector: target cell ratio of 25 : 1. △, Interferon alone; ○, interferon plus IL-2.

causing less stimulation. Nevertheless, even at those doses, the addition of IL-2 was associated with increased NK cell-mediated cytotoxicity, clearly implying that at least some of the potentiating effects of IL-2 differed from, and were additive to, those induced by interferon. This conclusion was supported by observations that several IL-2 preparations which had no demonstrable interferon activity, induced considerable augmentation of the cytotoxic activity of NK cells.

Table 1 illustrates the NK-cell potentiating effects of several IL-2 preparations as tested on BALB/c (nu^+/nu^+) responder spleen cells. Similar results were also found for CBA/J splenocytes. Concanavalin A (Con A)-induced rat IL-2 caused significant augmentation of murine NK cell activity, as did supernatants obtained from murine mixed lymphocyte cultures (MLCs). It is important to note that NK-cell potentiation was obtained using supernatants collected from a number of MLC combinations, including those (M locus differences) which do not elicit interferon production[12]. In agreement with the notion that IL-2 can potentiate NK-cell reactivity, the most potent augmentation of cytotoxicity was obtained with a purified IL-2 preparation from phytohaemagglutinin (PHA)-stimulated LBRM-33 murine lymphoma cells and purified by successive gel filtration, ion-exchange chromatography, preparative flat-bed isoelectric focusing and polyacrylamide gel electrophoresis (referred to in Table 1 as PAGE-purified IL-2; see ref. 13). IL-2 prepared from mitogen-stimulated LBRM cells by isoelectric focusing[14] was also active in potentiating NK-cell reactivity. Neither the isoelectrically purified IL-2 preparations nor those isolated by polyacrylamide gel electrophoresis showed interferon activity as assayed by virus neutralization[15] (M. Proffitt, personal communication). These findings strongly suggested that IL-2 was capable of augmenting NK cell-mediated cytotoxicity. Such conclusions were strengthened by studies in which IL-2 activity was specifically removed from the preparations.

A number of cells behave, at least operationally, as though they bear receptors for IL-2 (ref. 16). They remove conventional IL-2 activity from media containing the lymphokine in short-term adsorption trials (at 4 or 37 °C), regardless of whether the

Table 2 Loss of ability of IL-2 preparations to augment NK-cell reactivity following adsorption with Con A 'blasts'

Source of IL-2	Adsorption of IL-2 with:	IL-2 activity U ml^{-1}	NK cell mediated lysis
None	—	0	7.5
Rat spleen	None	1.25	36.8
	Normal spleen	1.10	30.6
	Con A 'blasts'	<0.05	5.4
None	—	—	12.3
Mouse MLC	None	2.3	32.5
	Normal spleen	2.1	33.8
	Con A 'blasts'	0.25	16.6

The NK-cell reactivity of CBA/J spleen cells was assessed after culture (16 h 37 °C) with or without, the IL-2 preparations noted. NK-cell mediated lysis was assessed in a 4-h ^{51}Cr release assay using YAC-1 target cells. The data shown were obtained at an effector : target cell multiplicity of 50:1. The IL-2 preparations used were obtained either from culture supernatants following Con-A stimulation of rat spleen (5 μg ml^{-1}, 24 h) or by collecting supernatants from 24-h MLC cultures between 10^7 B10.BR and 10^7 MLC-treated AKR spleen cells. In the case of mitogen-induced IL-2, a 1 : 40 dilution (1.25 U) was used, in the case of the MLC supernatant a 1 : 2 dilution (2.3 U) was used. The adsorption procedure was as follows: 100 μl of IL-2 containing medium was incubated (4 °C, 4 h) with (1) 10^7 normal C57BL/6 spleen cells or with (2) 10^7 C57BL/6 spleen cells previously stimulated for 24 h with 5 μg ml^{-1} Con A. At the end of incubation, the cells were removed by centrifugation and the supernatants filter-sterilized and tested for both NK augmenting activity and for conventional IL-2 activity.

adsorbing cell population is living, or fixed with glutaraldehyde[16]. Cells exhibiting this property include all IL-2-dependent effector T-cell lines thus far analysed, as well as mitogen-activated T-cell blasts[17]. Unstimulated lymphoid cells do not adsorb conventional IL-2 activity and thus apparently lack, or express very few, IL-2 receptors[16]. Culture supernatants from MLC-stimulated murine spleen cells, or Con A-stimulated rat spleen cells, were tested for their capacity to potentiate NK-cell reactivity both before and after adsorption (4 h, 4 °C) with either unstimulated or mitogen-activated T-cell populations (10^8 cells per ml). As can be seen in Table 2, the ability of the supernatants to potentiate the cytotoxic activity of NK cells was unaffected after adsorption with normal spleen cells. Similarly, adsorption with resting spleen cells had no effect on the ability of the supernant to induce the proliferation of an effector T-cell line (measured as U ml^{-1} IL-2 activity). The NK potentiating activity, as well as its conventional IL-2 activity, was however totally removed when Con A-activated splenocytes were used for adsorption. Both measurable T-cell proliferation and NK potentiation activities were similarly removed from the supernatants after adsorption with either of two long-term IL-2-dependent T-cell lines (data not shown).

We have recently described the preparation of a murine hybridoma which secretes an IgG antibody capable of inhibiting IL-2 activity[18]. This hybridoma was obtained by the fusion of spleen cells from BALB/c mice, previously immunized with rat IL-2, with the BALB/c myeloma SP2, according to the procedure of Kohler and Milstein[19]. We have isolated three hybridoma clones which secreted antibody that was able to neutralize (or precipitate, together with an additional complexing reagent) rat, mouse or human IL-2 activity. In tests of the effects of antibody with this specificity on the augmentation of NK-cell activity by IL-2 preparations (Table 3), although a small amount of NK-cell potentiating activity was lost nonspecifically by Protein-A bead mixing, only the anti-IL-2 antibody preparation was able to abrogate the NK-potentiating activity of the IL-2 preparations. This implies that the lymphokine IL-2 was, in fact, responsible for the augmentation of cytotoxicity by the lymphokine preparation.

There has been much recent interest in the ability of interferon (and of interferon-inducing agents) to augment NK-cell reactivity both *in vivo* and *in vitro*. Indeed, such observations seem to form the conceptual basis from which widespread clinical trials to assess the anti-neoplastic effects of interferon, have been launched. Thus far, the NK-augmenting effects of such agents as BCG, *Corynebacterium parvum* and poly I:C has been attributed exclusively to interferon. Although this may well be the case, this report indicates that interferon is not alone in its ability to promote NK cell-mediated cytotoxicity, leading one to question whether some of the potentiating affects previously attributed to interferon might be related not to this lymphokine, but to interleukin-2. In any case, our study suggests that IL-2, which has previously been shown to regulate a wide variety of immune responses, might be studied more closely as a potential immune response modifier, particularly in those situations in which NK cells may play a salient role.

This work was supported by NIH grants AI 15384, CA 24537 and CA 28419, grant IM-274 from the American Cancer Society and by a grant from the National Foundation. S.G. is a Special Fellow of the Leukemia Society of America. We thank Dr M. Proffitt for carrying out interferon assays on purified IL-2 preparations.

Table 3 Effect of anti-IL-2 monoclonal antibody on the potentiation of NK cell reactivity by IL-2 preparations

Effector Cell	IL-2 preparation from:	Antibody added	Passage through Protein-A-Sepharose column	IL-2 remaining after treatment (U ml^{-1}) Expt 1	IL-2 remaining after treatment (U ml^{-1}) Expt 2	NK-cell mediated lysis Expt 1	NK-cell mediated lysis Expt 2
BALB/c (nu^+/nu^+)	None	—	—	—	—	8.2	1.5
	Rat spleen	—	—	1.9	1.8	34.9	22.0
		—	+	1.5	1.7	24.0	18.9
		Anti-IL-2	+	0.2	<0.05	8.6	1.6
		Anti-gp70	+	ND	1.7	ND	18.1
CBA/J	None	—	—	—	—	10.0	7.6
	Rat spleen	—	—	3.8	1.9	38.5	22.8
		—	+	3.0	1.7	24.7	15.3
		Anti-IL-2	+	0.4	<0.05	6.8	1.0
		Anti-gp70	+	ND	1.6	ND	20.7

Cell-free supernatants obtained from Con A-stimulated rat splenocytes or 24-h murine MLC cultures (Table 1) were incubated with an equal volume of a monoclonal IgG antibody of indicated specificity (or with plain medium) for 30 min at 37 °C. These mixtures were then further incubated (or not) with protein A-conjugated Sepharose beads in an end-over-end shaker for 2 h at room temperature. The beads were then removed by centrifugation and the treated supernatants assessed for their ability to augment NK mediated cytolysis (of YAC-1 targets after culture for 16 h with BALB/c (nu^+/nu^+) or CBA/J spleen cells in a 4-h ^{51}Cr release assay) and concomitantly for conventional IL-2 activity (as measured (U ml^{-1}) by proliferation of an IL-2 dependent cell line). Con A (0.3 μg ml^{-1}) and supernatants from 24-h syngeneic MLC cultures were tested similarily at the same time for both NK augmenting activity and for conventional IL-2 activity and were found to give only background levels of each activity (data not shown).

Received 12 December 1980; accepted 10 March 1981.

1. Gillis, S. & Smith, K. A. *Nature* **268,** 154 (1977).
2. Nabholz, M., Engers, H., Collavo, D. & North, M. *Curr. Topics Microbiol. Immun.* **81,** 176 (1978).
3. Gillis, S., Baker, P. E., Ruscetti, F. W. & Smith, K. A. *J. exp. Med.* **148,** 1093 (1978).
4. Watson, J. *J. exp. Med.* **150,** 1510 (1979).
5. Gillis, S., Smith, K. A. & Watson, J. D. *J. Immun.* **124,** 1954 (1980).
6. Watson, J., Aarden, L. & Lefkovitz, I. *J. Immun.* **122,** 209 (1979).
7. Herberman, R. B., Nunn, M. E. & Lavrin, D. H. *Int. J. Cancer* **16,** 216 (1975).
8. Gillis, S., Ferm, M. M., Ou, W. & Smith, K. A. *J. Immun.* **120,** 2027 (1978).
9. Tracey, D. E., Wolfe, S. A., Durdik, J. M. & Henney, C. S. *J. Immun.* **119,** 1145 (1977).
10. Gidlund, M., Orn, A., Wigzell, H., Senik, A. & Gresser, I. *Nature* **273,** 759 (1978).
11. Djeu, J. Y., Heinbaugh, J. A., Holden, H. T. & Herberman, R. B. *J. Immun.* **122,** 175 (1979).
12. Kirchner, H., Zawatszky, R. & Schirrmacher, V. *Eur. J. Immun.* **9,** 97 (1979).
13. Watson, J., Mochizuki, D. M., Frank, M. B. & Gillis, S. in *Pharmacology of the Reticuloendothelial System* (ed. Webb, D.) (Dekker, New York, in the press).
14. Watson, J., Gillis, S., Marbrook, J., Mochizuki, D. & Smith, K. A. *J. exp. Med.* **150,** 849 (1979).
15. Hirsch, M. S., Proffitt, M. R. & Black, P. H. *Contemporary Topics Immunobiol.* **6,** 209 (1977).
16. Smith, K. A., Gillis, S., Ruscetti, F., Baker, P. E. & McKenzie, D. *Proc. N.Y. Acad. Sci.* **332,** 423 (1979).
17. Baker, P., Gillis, S. & Smith, K. A. *J. exp. Med.* **149,** 273 (1979).
18. Gillis, S. & Henney, C. S. *J. Immun.* (in the press).
19. Kohler, G. & Milstein, C. *Nature* **256,** 495 (1975).
20. Durdik, J. M., Beck, B. N., Clark, E. A. & Henney, C. S. *J. Immun.* **125,** 683 (1980).
21. Young, W. W., Hakomori, S., Durdik, J. M. & Henney, C. S. *J. Immun.* **124,** 199 (1980).
22. Ledbetter, J. & Nowinski, R. C. *J. Virol.* **23,** 315 (1977).

Synthesis and secretion of mouse immunoglobulin chains from *Xenopus* oocytes

Giorgio Valle, Jayne Besley & Alan Colman

Department of Biological Sciences, University of Warwick,
Coventry CV4 7AL, UK

The secretion of immunoglobulins from mouse plasmacytoma cell lines is a valuable model for the study of mechanisms of protein secretion from eukaryotic cells. Some cells do not synthesize heavy chain immunoglobulin while retaining the ability to synthesize though not to secrete the complementary light chain[1–3]. In the MOPC 21 variant line, NS1, it is absence of heavy chain that prevents secretion[4–6]. A reciprocal dependence, that is, a requirement of light chain for heavy chain secretion, has not been tested directly because no variants producing normal heavy chain alone have been detected[4,7–9]. In view of the known insolubility characteristics of isolated heavy chains *in vitro*[10–12], this failure has led to the idea that free heavy chains would be toxic to the cell[7,9]. Recently, *Xenopus laevis* oocytes have been shown to provide a surrogate system for studying protein secretion after microinjection of the encoding mRNA[13,14]. Using this technique we demonstrate here that tetrameric mouse immunoglobulin is assembled and secreted by the oocyte but that free light chains are not secreted. Moreover, using purified heavy chain mRNA we show that heavy chain dimers accumulate in the oocyte and are not secreted unless light chain mRNA is also injected.

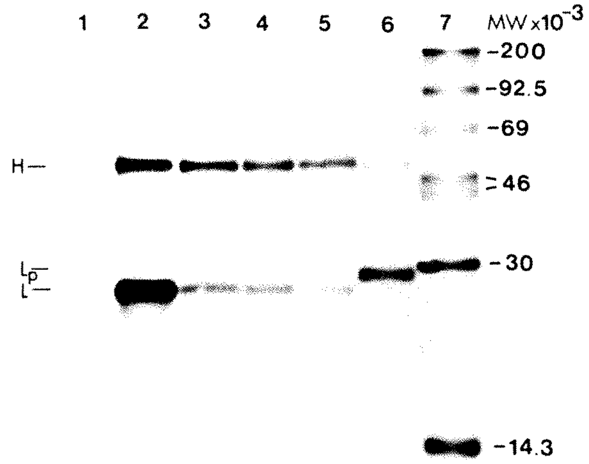


Fig. 1 Synthesis and secretion of X63 immunoglobulin. Culture and labelling of plasmacytoma cells: P3/X63, P3/NS1 and PuBu1 cells (supplied by D. Secher) were grown in Dulbecco's modified Eagle's medium containing 10% horse serum (Difco). Cells were labelled with ^{35}S-methionine (Radiochemical Centre) as described by Cotton *et al.*[2]. Media and cell homogenates were retained for immunoprecipitation and electrophoresis. mRNA preparation: 5×10^9 cells were lysed in hypotonic buffer[21] containing 0.1% diethylpyrocarbonate (Sigma). The post-nuclear supernatant was extracted with phenol and then precipitated with ethanol. The RNA pellet was resuspended in 5 ml of 10 mM Tris-HCl *p*H 7.5, 1 mM EDTA, 0.1% SDS and 0.5 M LiCl, bound to oligo(dT)-cellulose (type III, Collaborative Research) and eluted with the same buffer minus LiCl. The polyadenylated mRNA was then ethanol precipitated and resuspended in distilled water at 1 mg ml^{-1}. Microinjection and culture of oocytes: oocytes were microinjected with 30 nl of mRNA and cultured in Barths saline (5 oocytes per 30 µl) containing 1 µCi µl^{-1} ^{35}S-methionine (300 Ci mmol^{-1}) for 24 h at 21 °C (ref. 13). Culture media were retained for electrophoresis and oocytes were fractionated on sucrose step gradients as previously described[13]. Essentially two fractions were obtained: a vesicle fraction derived from oocyte intracellular membranes and a cytosol fraction containing the soluble components of the oocyte. Immunoprecipitation and electrophoresis: 50-µl aliquots of cell lysates, oocyte fractions and culture media were immunoprecipitated with 3 µl of goat anti-mouse IgG (Miles) using *Staphylococcus aureus* envelopes as described by Curtis *et al.*[22], the only modification being the use of immunoprecipitation buffer containing 1% Triton X-100 (v/v), 1% sodium deoxycholate (w/v), 0.5% SDS (w/v), 1 mM phenylmethylsulphonyl fluoride, 100 mM KCl, 5 mM $MgCl_2$ and 100 mM Tris-HCl, *p*H 8.2 (A. Williamson, personal communications). Precipitates were dissolved in sample buffer (20 mM Tris-HCl *p*H 7.6, 20% glycerol, 1 mM phenylmethylsulphonyl fluoride, 1% β-mercaptoethanol, 0.01% bromophenol blue and 2% SDS, and 25-µl aliquots run on 12.5% polyacrylamide gels[23], followed by fluorography[24]. Track 1, oocyte cytosol; 2, oocyte vesicle; 3, oocyte medium; 4, X63 cell lysate; 5, X63 cell medium; 6, X63 mRNA translated in reticulocyte lysate[15]; 7, molecular weight (MW) markers (Radiochemical Centre). Abbreviations: H, heavy chain; L, light chain; L_p, light chain produced in reticulocyte lysate.

The cell line P3/X63 produces an immunoglobulin (MOPC 21) consisting of γ1 heavy chains and κ light chains[15]. When mRNA from this cell line is injected into oocytes both heavy (H) and light (L) chain proteins are synthesized, sequestered within membranes and secreted (Fig. 1, tracks 1–3); these chains co-migrate with the analogous products prepared from the X63 cells and media (Fig. 1, tracks 4, 5). However, comparison of the ^{35}S-methionine contents of secreted and intracellular heavy and light chains reveals that a disproportionate amount of heavy chain is secreted by oocytes, though not by cultured cells. In fact, when allowance is made for the different numbers of methionine residues in each chain (four in light chain, nine in heavy chain[16,17]), it is evident that whereas approximately equal amounts of each chain are secreted, excess light chain accumulates within the oocytes. As only assembled (tetrameric H_2L_2) IgG is secreted from the X63 cells[5], these results suggest that only assembled IgG is secreted by the oocyte. This possibility was examined by analysing the intra- and extracellular immunoglobulin on non-reducing polyacrylamide gels. As can be seen in Fig. 2*a* (track 2) two major species are present intracellularly in the oocyte, corresponding to the tetrameric IgG and free light chains. Surprisingly, very little tetramer can be seen in the oocyte medium (Fig. 2*a*, track 3), in contrast to the X63 medium (track 5). Instead, trimers, dimers and monomers are observed. However, when 6% dialysed horse serum is included in the oocyte medium (Fig. 2*b*), most of the secreted IgG is in the tetrameric form (compare tracks 2 and 4); presumably, the added serum reduces gratuitous oxidation of the external immunoglobulin.

A further demonstration of the fidelity of the oocyte in its surrogate role is seen after microinjection of the mRNAs encoding the κ light chains of the cell lines NS1 and P1. Of these two light chains only the P1 κ-chain is secreted *in vivo* both in an assembled form (with the P1 γ2a heavy chain) and free[5]. When P1 and NS1 mRNAs are injected separately or together, it is clear that only the P1 κ-chain is secreted even when both κ-chains co-exist in the same oocyte (Fig. 3). Although a rescue of the NS1 κ-chain by assembly with the P1 γ2a heavy chain might have been anticipated, in view of a similar phenomenon in P1/NS1 hybrids[5], very little γ2a heavy chain was synthesized in

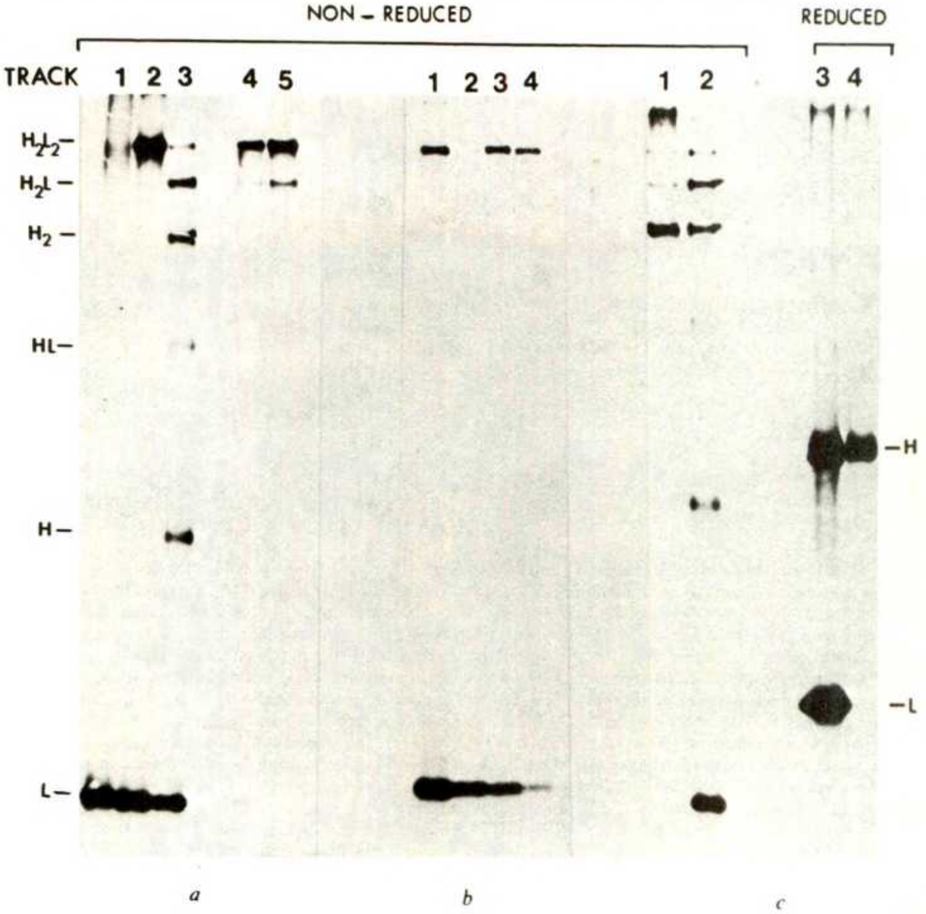


Fig. 2 Assembly of immunoglobulins in oocytes. Oocytes were injected with X63 mRNA, NS1 mRNA or X63 heavy chain mRNA (see Fig. 4*a* legend), cultured and fractionated as described in Fig. 1 legend. Aliquots (50-μl) were then alkylated with iodoacetic acid before immunoprecipitation. Precipitates were eluted from *S. aureus* in 50 μl of sample buffer without β-mercaptoethanol, boiled at 100 °C for 2 min and divided into two 25-μl aliquots. One aliquot (non-reduced) was ready for use, the second was reduced and completely alkylated[25] before electrophoresis on exponential 7.5–20% gradient gels as described in Fig. 1 legend; *a–c* illustrate independent experiments. *a*, Track 1, oocyte vesicles, NS1 mRNA; 2, oocyte vesicles, X63 mRNA; 3, oocyte medium, X63 mRNA; 4, X63 cell lysate; 5, X63 cell medium. *b*, Track 1, oocyte vesicles, X63 mRNA; 2, oocyte medium, X63 mRNA; 3 and 4, as tracks 1 and 2, respectively, except that incubation medium contained 6% dialysed horse serum. *c*, Track 1, oocyte vesicles, heavy chain mRNA; 2, oocyte medium, X63 mRNA; 3, oocyte vesicles, X63 mRNA; 4, as track 1.

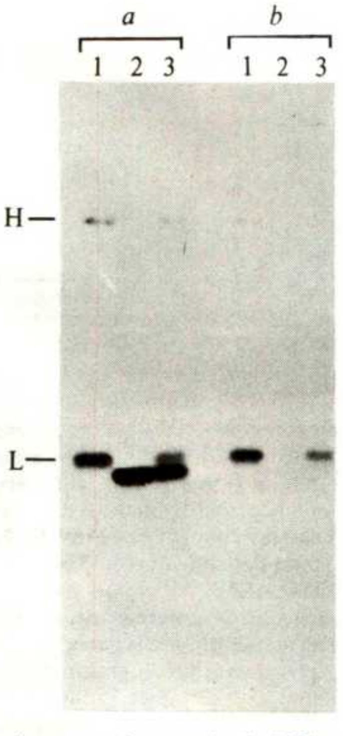


Fig. 3 Synthesis and secretion of NS1 and P1 immunoglobulins. Oocytes were microinjected with NS1 or P1 mRNAs or NS1 plus P1 mRNA and incubated in ^{35}S-methionine as described in Fig. 1 legend. Oocytes were then homogenized in phosphate-buffered saline (40 μl per oocyte) containing 1% Triton X-100 and 1 mM phenylmethylsulphonyl fluoride and then clarified by centrifugation. Oocyte homogenate and incubation media were immunoprecipitated, alkylated and then electrophoresed on 12.5% gels. In terms of amount per oocyte, six times more sample was electrophoresed in the oocyte media tracks than the corresponding homogenate tracks. Homogenates or media were prepared from oocyte injected with P1 mRNA (1), NS1 mRNA (2), NS1 + P1 mRNA (3). *a*, Oocyte homogenate; *b*, oocyte incubation media. H, heavy chain; L, light chain.

these experiments. As the NS1 and X63 light chains are identical in amino acid sequences (C. Milstein, personal communication), these experiments demonstrate that the NS1 light chain cannot be secreted from oocytes in the absence of heavy chain (compare Fig. 3, tracks 2, with Fig. 1, tracks 2, 3).

As mentioned earlier, cell lines lacking light chain but synthesizing normal γ heavy chain have not been detected. We have tested the intrinsic secretory potential of isolated heavy chains by separating heavy and light chain mRNAs on formamide–agarose gels (Fig. 4*a*) and then injecting the recovered mRNAs both separately and recombined into oocytes. As can be seen in Fig. 4*b*, heavy chain alone is synthesized and sequestered within oocyte membranes but it is not secreted unless purified light chain mRNA is also injected. This rescue is also affected by co-injection of either NS1 mRNA (Fig. 4*b*) or the MOPC 315 λII light chain mRNA (A. Colman, C. Valle, T. R. Mosmann and A. R. Williamson, data not shown). Further experiments involving various permutations of heavy and light chain mRNA fractions show that the amount of heavy chain rescued depends, in conditions of light chain excess (Fig. 4*c*), on the amount of heavy chain mRNA present in the oocyte. This argues against the stoichiometric assistance of some putative plasmacytoma helper protein in immunoglobulin secretion unless the respective mRNA exactly co-migrates with heavy or light chain mRNAs. Our data indicate that the X63 heavy chain is intrinsically non-secretory unless light chain is present. Such a situation occurs in variant lines which synthesize abnormal heavy chain and no light chain; the heavy chains are not secreted unless light chains are re-introduced through cell fusion methods[9]. What is the reason for this non-secretory phenotype? It might be argued that as a result of its known *in vitro* insolubility[10–12], intracellular heavy chain would aggregate in the absence of light chain and then be unavailable for secretion. The results shown in Fig. 2*c* show that a substantial proportion of the stranded heavy chain is in the form of covalent dimers and not the covalent aggregates which have been observed *in vitro*.

We conclude that the oocyte both assembles and secretes immunoglobulin; the assembly of IgG in oocytes has been noted by others[18]. If this immunoglobulin retains its antibody specificity, as seems likely[19], the contribution of the individual chains to this specificity could be assessed by generating different combinations of heavy and light chain by means of mRNA injection. We do not understand why the separated heavy and light MOPC 21 chains are not secreted. The evidence presented above indicates that the chains exist in oocytes as monomers (light chain) or dimers (heavy chain) and we have preliminary evidence that they do not form strong attachments

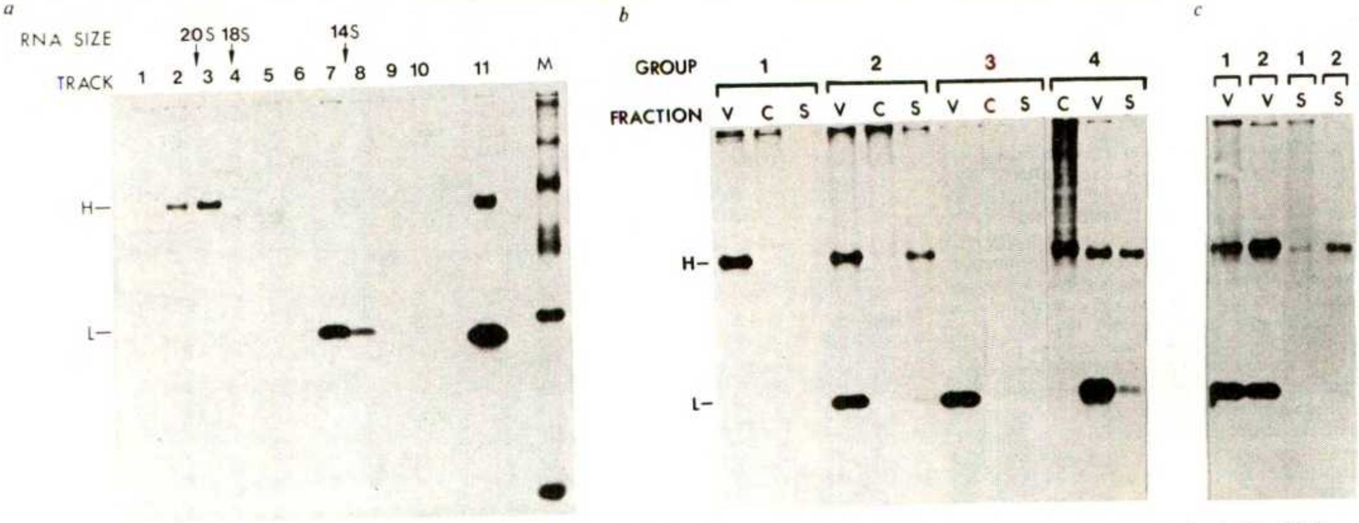

a
RNA SIZE
20S 18S
14S
TRACK 1 2 3 4 5 6 7 8 9 10 11 M
H—
L—
b
GROUP 1 2 3 4
FRACTION V C S V C S V C S C V S
H—
L—
c
1 2 1 2
V V S S

Fig. 4 Microinjection of separated heavy and light chain mRNAs. *a*, 30 μg of poly(A) plus RNA (Fig. 1) were fractionated by electrophoresis (40 mA, 5 h, 4 °C) in 1.5% agarose gel (Sigma, type II, medium EEO) made in electrophoresis buffer (50% formamide, 0.1 mM EDTA, 3.6 mM Tris and 4 mM NaH_2PO_4, *p*H 7.4); gel dimensions were 0.3 × 16 × 16 cm. As MW marker, 5 μg of *Xenopus laevis* total ovary RNA was also run on a neighbouring 8-mm wide slot. The part of the gel containing the X63 RNA was cut into 5-mm slices. Each slice (1 cm^3) was squeezed through a 23G needle into 1 ml of 10 mM Tris, 1 mM EDTA and 0.1% SDS, *p*H 7.5, and shaken at 4 °C for ~20 h. After removal of polymeric agarose by centrifugation at $180{,}000g_{max}$ for 30 min at 4 °C, poly(A)$^-$ RNA from *X. laevis* ovary (a gift of P. C. Turner) was then added to each fraction as carrier. After ethanol precipitation the RNA pellet was resuspended in 30 μl of distilled water and spun for 3 min in an Eppendorf microfuge before microinjection. Batches of oocytes were injected with carrier RNA (track 1), mRNA samples corresponding to consecutive fractions of the gel (tracks 2–10) or the original X63 mRNA (track 11) and then incubated for 24 h in ^{35}S-methionine. Alkylated immunoprecipitates prepared from oocyte homogenates were electrophoresed as described in Fig. 1 legend. The arrows at the top of the figures indicate the approximate size of the RNA injected as deduced from visualization with ethidium bromide of the track on the RNA gel containing the RNA marker. *b*, *c*, Oocytes were injected with fractionated or non-fractionated mRNAs as indicated below and cultured as described in Fig. 1 legend, then vesicle (V), cytosol (C) and incubation medium (S) fractions were immunoprecipitated, alkylated and electrophoresed on 12.5% polyacrylamide gels; *b* and *c* illustrate independent experiments. *b*, Group 1, heavy chain mRNA; 2, heavy chain plus light chain mRNAs; 3, light chain mRNA; 4, NS1 mRNA + heavy chain mRNA. *c*, Diluted (track 1) or undiluted (track 2) heavy chain mRNA + light chain mRNA. H, heavy chain; L, light chain.

to membranes. It is therefore possible that the non-appearance of plasmacytoma variants synthesizing only heavy chain is not caused by the aggregation of this protein. Thus, the rescue of heavy chain by light chain in oocytes might not be related to the reported mitigating effect of light chain on heavy chain aggregation *in vitro*[12].

As the H_2 dimer is known to be an intermediate form in the assembly of this subclass of immunoglobulins both *in vivo*[20] and *in vitro*[12], it will be interesting to see whether stranded H_2 dimers can be rescued by the subsequent injection of light chain mRNA. Such a demonstration would indicate that the topological block to secretion of isolated light chain occurs at the same point or further along the secretory pathway than the block to heavy chain secretion.

We thank Drs A. Williamson, C. Milstein and D. S. Secher for useful discussion, the Cancer Research Campaign for support, and S. Bhamra for technical assistance.

Received 4 December 1980; accepted 2 March 1981.

1. Coffino, P. & Scharff, M. D. *Proc. natn. Acad. Sci. U.S.A.* **68,** 219–223 (1971).
2. Cotton, R. G. H., Secher, D. S. & Milstein, C. *Eur. J. Immun.* **3,** 135–140 (1973).
3. Mosmann, T. R., Baumal, R. & Williamson, A. R. *Eur. J. Immun.* **9,** 511–516 (1979).
4. Cowan, N. J., Secher, D. S. & Milstein, C. *J. molec. Biol.* **90,** 691–701 (1974).
5. Kohler, G., Howe, S. C. & Milstein, C. *Eur. J. Immun.* **6,** 292–295 (1976).
6. Milstein, C. *et al. Cold Spring Harb. Symp. quant. Biol.* **41,** 793–803 (1977).
7. Kohler, G. *Proc. natn. Acad. Sci. U.S.A.* **77,** 2197–2199 (1980).
8. Morrison, S. L. & Scharff, M. D. *J. Immun.* **114,** 655–659 (1975).
9. Wilde, C. & Milstein, C. *Eur. J. Immun.* **10,** 462–469 (1980).
10. Bjork, I. & Tanford, C. *Biochemistry* **10,** 1271–1280 (1971).
11. Stevenson, G. T. & Dorrington, K. J. *Biochem. J.* **118,** 703–710 (1970).
12. Sears, D. W., Kazin, A. R., Mohrer, J., Friedman, F. & Beychok, S. *Biochemistry* **16,** 2016–2025 (1977).
13. Colman, A. & Morser, J. *Cell* **17,** 517–526 (1979).
14. Lane, C. D. *et al. Eur. J. Biochem.* **111,** 225–235 (1980).
16. Svasti, J. & Milstein, C. *Biochem. J.* **128,** 427–444 (1972).
17. Adetugbo, K., Milstein, C. & Secher, D. S. *Nature* **265,** 299–304 (1977).
18. Stevens, R. H. & Williamson, A. R. *J. molec. Biol.* **78,** 517–525 (1973).
19. Deacon, N. J. & Ebringer, A. *Biochem. Soc. Trans.* **4,** 818–820 (1976).
20. Bevan, M. J., Parkhouse, R. M. E., Williamson, A. R. & Askonas, B. A. *Prog. Biophys. molec. Biol.* **25,** 131–162 (1972).
21. Harrison, T. H., Brownlee, G. G. & Milstein, C. *Eur. J. Biochem.* **47,** 613–620 (1974).
22. Curtis, P., Mantei, N., Van der Berg, J. & Weissman, C. *Proc. natn. Acad. Sci. U.S.A.* **75,** 1309–1313 (1978).
23. Laemmli, U. K. *Nature* **227,** 680–685 (1970).
24. Bonner, W. & Laskey, R. *Eur. J. Biochem.* **46,** 83–88 (1974).
24. Pelham, H. R. B. & Jackson, R. J. *Eur. J. Biochem.* **67,** 247–256 (1976).
25. Mosmann, T. R., Gravel, Y., Williamson, A. R. & Baumal, R. *Eur. J. Immun.* **8,** 94–101 (1978).

Transcriptional regulation of mouse liver metallothionein-I gene by glucocorticoids

Lisa J. Hager & Richard D. Palmiter

Howard Hughes Medical Institute Laboratory,
Department of Biochemistry, University of Washington, Seattle,
Washington 98195, USA

Glucocorticoid hormones play an important part in the regulation of many essential metabolic processes[1]. These include the synthesis of stress-related 'acute-phase' proteins[2] and the timing of developmental events in numerous tissues including fetal lung, liver and intestine, and adult mammary gland[1]. Among the proteins that are induced by glucocorticoids during stress and fetal development are the metallothioneins (MT)[3–6]. These small, cysteine-rich proteins are found primarily in the liver and kidney. They bind heavy metals and are thought to be involved in zinc homeostasis and heavy metal detoxification[7]. We show here that the accumulation of MT mRNA in mouse liver in response to the synthetic glucocorticoid, dexamethasone, is due to transcriptional activation of the metallothionein gene.

We began our study of MT expression *in vivo* by examining the kinetics of MT mRNA induction with dexamethasone. The probe used to quantitate mRNA levels was prepared by nick-translation of a cDNA clone containing the entire coding region of metallothionein-I (MT-I)[8]. The cDNA strand was separated from the message strand using a single-strand phage containing the MT-I message strand[9]. Initial experiments indicated that the kinetics of the MT-I mRNA induction were influenced by both the dose and the preparation of dexamethasone. A soluble form of dexamethasone elicited a rapid response, such that hepatic MT-I mRNA levels were maximal at 2 h and declined to basal levels by 6 h. With an insoluble, 'slow release' dexamethasone, MT-I mRNA levels rose gradually for 6–9 h. To obtain a MT-I

0028-0836/81/220340—03$01.00

Table 1 Correlation between mRNA accumulation, transcription and the presence of specific glucocorticoid receptors

		Nuclear radioactivity (c.p.m. per μg DNA)		Nuclear receptors (molecules per cell)	MT-I mRNA (molecules per cell)	MT-I gene transcription (p.p.m.)
		50 μg Dex	5 mg Dex			
Liver	Control	—	—	—	13	3
	Dex	10.4 ∓ 1.49*	1.23 ∓ 0.24	59,500	1,360	260
Kidney	Control	—	—	—	14	4
	Dex	1.98 ∓ 0.56	0.61 ∓ 0.29	8,800	39	13

Six mice were each injected with ^{3}H-dexamethasone (Dex, 100 μCi; 84 Ci mmol^{-1}; NEN) and 50 μg of the soluble form of dexamethasone. Three of the mice also received a 100-fold excess (5 mg) of the soluble dexamethasone to monitor nonspecific nuclear binding of the ^{3}H-dexamethasone. After 2 h, liver and kidney nuclei were isolated and aliquots (400–1,000 μg of DNA) were counted at 35% efficiency. The DNA content of each sample was determined by a diphenylamine assay[32]. Receptor molecules per cell were calculated by subtracting column *b* from column *a* and using values of 770 mCi mmol^{-1}, 35% counting efficiency and 6.4 pg of DNA per cell. The mRNA levels were measured after 6 h of hormone treatment. The relative rate of transcription was determined after 2 h of hormone treatment (Fig. 3).

* Standard deviation ($n = 6$).

mRNA induction that was both rapid and prolonged, we used equal amounts of both seroid forms.

Figure 1 shows a composite dose response, based on several experiments using male and female mice of the Swiss Webster and BALB/c strains. Maximal induction of MT-I mRNA, measured at 6 h, was achieved with > 70 μg of dexamethasone.

The kinetics of MT-I mRNA induction with an optimal dose of dexamethasone (100 μg) are shown in Fig. 2. The MT-I mRNA level increased ~100-fold to about 1,350 molecules per cell by 6 h and then gradually fell, presumably due to depletion of the steroid. We detected little difference in the kinetics or extent of mRNA accumulation between male and female mice in either mouse strain. Occasionally we observed a high basal level of MT-I mRNA in liver, sometimes as high as 300 molecules per cell, which is probably a stress phenomenon related to handling or cage conditions.

Steroids are generally assumed to act at the transcriptional level, although effects on mRNA stability have also been reported[10,11]. Several different mRNAs are induced by glucocorticoids, including those coding for tyrosine aminotransferase[12], tryptophan oxygenase[13], ovalbumin[14], transferrin[15], growth hormone[16], $\alpha_{2\mu}$-microglobulin[17] and metallothionein[18,19]. However, the only direct demonstrations of transcriptional control by glucocorticoids are for mammary tumour virus expression[20,21] and possibly casein gene expression[22]. To ascertain whether MT-I mRNA accumulation is a consequence of transcriptional activation, liver nuclei were isolated at various times after hormone administration and the endogenous RNA polymerases were allowed to transcribe a few hundred nucleotides in the presence of [α- ^{32}P]UTP. The ^{32}P-RNA was hybridized to immobilized plasmid DNA containing the genomic MT-I sequence. This assay essentially measures the relative number of RNA polymerases on the MT-I gene at the time the nuclei are isolated[11]. The histograms in Fig. 2 show that the relative rate of transcription of the MT-I gene increased ~90-fold within the first 2 h and stayed high for another 7 h. Similar effects of dexamethasone on MT mRNA synthesis are observed with cultured mouse cells[23]. The kinetics of transcriptional activation predict the mRNA levels reasonably well,

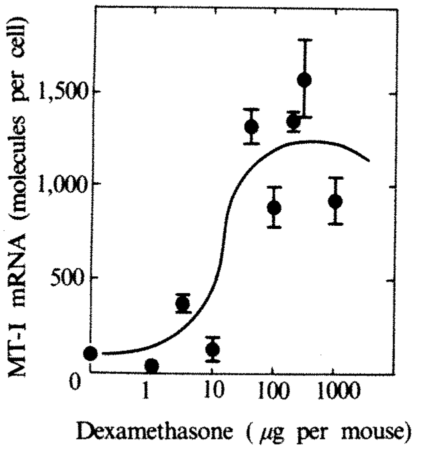


Fig. 1 Dose response of MT-I mRNA induction by dexamethasone. Mice (23–27 g) were injected subcutaneously with the indicated amount of dexamethasone or saline (0.2 ml). Fifty per cent of the dose was dexamethasone phosphate (Decadron; Merck, Sharp and Dohme) and 50% was dexamethasone acetate (Decadron-LA; Merck, Sharp and Dohme). At 6 h, the mice were killed and the livers were used to isolate total nucleic acid for mRNA measurements[9,11]. MT-I mRNA levels were determined by hybridizing total nucleic acid samples to a ^{32}P-cDNA probe prepared by nick-translation of a cDNA clone which contains the entire coding region and 3′-untranslated region of MT-I[8]. The extent of hybridization was determined by S_1 nuclease resistance and compared with a standard curve generated from the hybridization of the probe to known amounts of a recombinant bacteriophage fd DNA containing the message strand of MT-I[9]. Results are reported as mRNA molecules per cell, assuming a molecular weight of 130,000 for MT-I mRNA and that mouse liver cells contain 6.4 pg of DNA. Error bars show the standard error (n = 3–9) for each point.

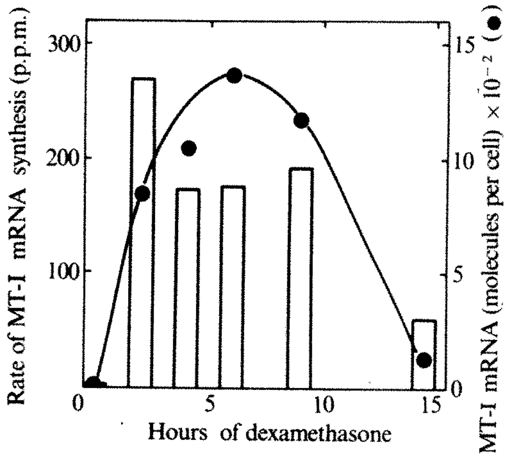


Fig. 2 Time course of MT-I mRNA accumulation and the relative rate of transcription in mouse liver. Tissue from mice induced for various times with 100 μg of dexamethasone was used to measure mRNA levels as in Fig. 1 (●). Nuclei from the same tissue samples were isolated and transcribed *in vitro* in the presence of [α-^{32}P]UTP[11]. ^{32}P-RNA transcripts were isolated and three different amounts of each sample were hybridized to plasmid DNA immobilized on nitrocellulose filters[11]. Each hybridization reaction contained a filter with 0.7 μg of $m_1pSH_{1.3}$ (a plasmid containing a 1.3-kilobase genomic DNA insert representing the entire MT-I gene[8,9]) and a control filter with 0.7 μg of pBR322. Relative rates of transcription (histograms) are reported as parts per million (p.p.m.) and were derived from the slope of the line generated by plotting radioactivity hybridized versus total input radioactivity (see Fig. 3). Background (< 10 p.p.m.) measured by the pBR322 filters, was subtracted. Values are corrected for 44% hybridization efficiency, as determined by the hybridization of MT-I mDNA in a parallel reaction.

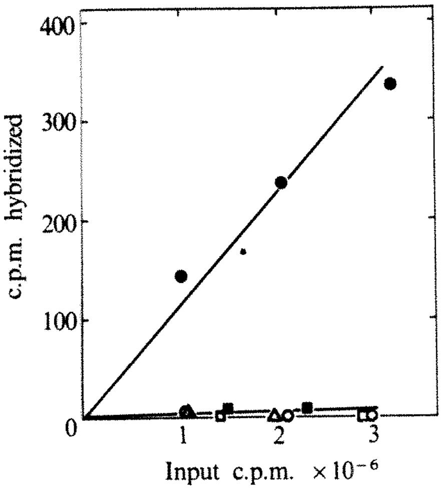


Fig. 3 Effect of α-amanitin on MT-I gene transcription. Liver nuclei from mice induced with dexamethasone for 2 h were transcribed in the presence of [α-^{32}P]UTP (●); duplicates included 5 μg ml^{-1} α-amanitin (△). Varying amounts of ^{32}P-RNA transcripts were hybridized as described in Fig. 2 legend. pBR322 controls have been subtracted. Also shown are the results of the transcription of uninduced liver nuclei (○), and kidney nuclei from induced (■) and uninduced (□) mice.

although small effects on mRNA stability cannot be entirely ruled out.

Figure 3 illustrates the effect of α-amanitin on the relative rate of MT-I gene transcription. When liver nuclei isolated 2 h after dexamethasone treatment were allowed to transcribe in the presence of α-amanitin, total transcription was inhibited by 54% and specific MT-I transcription dropped from 260 p.p.m. to less than 10 p.p.m. Thus, we conclude that MT-I mRNA is transcribed by RNA polymerase II (form B), as are all other mRNAs.

Heavy metals such as zinc and cadmium also induce mouse liver MT-I mRNA by transcriptional activation[24]. Maximal levels of transcription and mRNA accumulation (2,200 molecules per cell) are 1.5- to 2-fold higher than those obtained with dexamethasone; thus, there is a good correlation between the rate of transcription and mRNA accumulation with both metal and hormonal inducers. The ratio of the amount of MT-I protein that accumulates in the liver to the amount of MT-I mRNA induced by glucocorticoids or metals has not been determined adequately, primarily because of the difficulty in achieving steady-state conditions *in vivo*.

With dexamethasone, MT-I mRNA induction is most prominent in liver but we were also able to measure low levels in other tissues such as heart, skeletal muscle, spleen and kidney. MT-I mRNA accumulation and transcription are barely detectable in the kidney (Fig. 3, Table 1), whereas this tissue accumulates about 700 molecules of MT-I mRNA per cell in response to cadmium[24]. To explore the basis of this difference in responsiveness, we measured the number of nuclear glucocorticoid receptors in the liver and kidney. Table 1 shows that there are 6.5-fold more nuclear glucocorticoid receptors (59,500 molecules per cell) in the liver than in the kidney (8,800 molecules per cell). The number of nuclear glucocorticoid receptors found in the liver is consistent with the number of total cytoplasmic receptors ($\sim 10^5$ per cell) calculated from the data of Wrange *et al.*[25]. Assuming that mRNA induction is proportional to nuclear receptor levels[23], the paucity of glucocorticoid receptors accounts largely for the meagre response of the kidney relative to the liver. In addition, the data suggest that a lower percentage of the cells in the kidney than the liver may be responsive to either inducer. The minimal response of the kidney to glucocorticoids is consistent with our observations, as well as those of Oh *et al.*[3], that stressed animals have high levels of MT-I mRNA and protein in the liver but not in the kidney.

The finding that MT genes are under glucocorticoid control may be explained in several ways. The available data suggest that the apoproteins synthesized from MT mRNA bind zinc predominantly; thus, induction of these proteins may increase the intracellular pool of this essential metal[3,6,19,26]. Recent studies indicate that MT may transfer zinc directly to other zinc-requiring apoenzymes[27,28]. A requirement of zinc for rapid tissue growth may account for the high levels of fetal MT that occur just before parturition[4,5]. Notably, the induction of fetal MT coincides with a dramatic rise in circulating glucocorticoids[29,30]. Because many heavy metals will displace zinc from MT[31], the induction of zinc-MT by glucocorticoids could also provide insurance against subsequent exposure to other toxic metals.

This work was supported by NIH grant HD-09172.

Received 15 December 1980; accepted 12 March 1981.

1. Baxter, J. D. & Rousseau, G. G. (eds) *Monogr. Endocr.* **12** (1979).
2. Grieninger, G., Hertzberg, K. M. & Pindyck, J. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5506–5510 (1978).
3. Oh, S. H., Deagen, J. T., Whanger, P. D. & Weswig, P. H. *Am. J. Physiol.* **234**, E282–E285 (1978).
4. Ryden, L. & Deutsch, H. F. *J. biol. Chem.* **253**, 519–524 (1978).
5. Wong, K.-L. & Klaassen, C. D. *J. biol. Chem.* **254**, 12399–12403 (1979).
6. Karin, M. & Herschman, H. R. *Eur. J. Biochem.* **113**, 267–272 (1981).
7. Kägi, J. H. R. & Nordberg, M. (eds) *Metallothionein* (Birkhauser, Basel, 1979).
8. Durnam, D. M., Perrin, F., Gannon, F. & Palmiter, R. D. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6511–6515 (1980).
9. Beach, L. R. & Palmiter, R. D. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
10. Yamamoto, K. R. & Alberts, B. M. *A. Rev. Biochem.* **45**, 721–745 (1976).
11. McKnight, G. S. & Palmiter, R. D. *J. biol. Chem.* **254**, 9050–9058 (1979).
12. Roewekamp, W. G., Hofer, E. & Sekeris, C. E. *Eur. J. Biochem.* **70**, 259–268 (1976).
13. Delap, L. & Feigelson, P. *Biochem. biophys. Res. Commun.* **82**, 142–149 (1978).
14. Hager, L. J., McKnight, G. S. & Palmiter, R. D. *J. biol. Chem.* **255**, 7796–7800 (1980).
15. Lee, D. L., McKnight, G. S. & Palmiter, R. D. *J. biol. Chem.* **253**, 3494–3503 (1978).
16. Tushinski, R. J., Sussman, P. M., Yu, L.-Y. & Bancroft, F. C. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2357–2361 (1977).
17. Kurtz, D. T., Chan, K.-M. & Feigelson, P. *J. biol. Chem.* **253**, 7886–7890 (1978).
18. Karin, M. & Herschman, H. R. *Science* **204**, 176–177 (1979).
19. Etzel, K. R., Shapiro, S. G. & Cousins, R. J. *Biochem. biophys. Res. Commun.* **75**, 806–812 (1979).
20. Ringold, G. M., Yamamoto, K. R., Tomkins, G. M., Bishop, J. M. & Varmus, H. E. *Cell* **6**, 299–305 (1975).
21. Stallcup, M. R., Ring, J. & Yamamoto, K. R. *Biochemistry* **17**, 1515–1521 (1978).
22. Ganguly, R., Mehta, N. M., Ganguly, N. & Banerjee, M. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6466–6470 (1979).
23. Mayo, K. E. & Palmiter, R. D. *J. biol. Chem.* **256**, 2621–2624 (1981).
24. Durnam, D. M. & Palmiter, R. D. *J. biol. Chem.* (in the press).
25. Wrange, O., Carlstedt-Duke, J. & Gustafsson, J.-A. *J. biol. Chem* **254**, 9284–9290 (1979).
26. Failla, M. L. & Cousins, R. J. *Biochim. biophys. Acta* **543**, 293–304 (1978).
27. Udom, A. D. & Brady, F. O. *Biochem. J.* **187**, 329–335 (1980).
28. Li, T.-Y., Kraker, A. J., Shaw, C. F. & Petering, D. H. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6334–6338 (1980).
29. Barlow, S. M., Morrison, P. J. & Sullivan, F. M. *J. Endocr.* **60**, 473–483 (1974).
30. Dalle, M., Giry, J., Gay, M. & Delost, P. *J. Endocr.* **76**, 303–309 (1978).
31. Cherian, M. G. & Goyer, R. A. *Life Sci.* **23**, 1–10 (1978).
32. Burton, K. *Meth. Enzym.* **12**, 163–166 (1968).

Unilateral internuclear gene transfer and cell differentiation in *Schizophyllum*

Richard F. Gaber[*] & Thomas J. Leonard[†]

Departments of Botany and Bacteriology, University of
Wisconsin, Madison, Wisconsin 53706, USA

The recent proliferation of studies on moveable genetic elements has suggested that genome rearrangement can contribute to the underlying mechanisms involved in cell differentiation[1]. Evidence to support this notion is rapidly accumulating in both prokaryotic and eukaryotic organisms[2–5]. So far, however, relatively few eukaryotes have lent themselves to genetic analysis with resolution sufficient for the detection of genome rearrangement during cell differentiation. The filamentous fungi as exemplified here by *Schizophyllum commune*, provide an unusual experimental opportunity for detecting moveable genetic elements in that the organism grows as a binucleate heterokaryon. Each cell contains two compatible haploid nuclei

---

[*] Present address: Department of Genetics, University of Wisconsin, Madison, Wisconsin 53706, USA.

[†] To whom requests for reprints should be addressed.

which can be experimentally dissociated into uninucleate cells, and these cells can then be used in genetic analysis to detect genetic changes. We now present evidence that nuclei in cells from differentiated mound bodies show a specific genetic change whereas nuclei from surrounding non-mound vegetative cells show no detectable genetic changes. The binucleate cells of differentiated mound bodies were found to contain haploid nuclei one of which remained unchanged with regard to its genetic marker genes whereas the other nucleus, and always the same one, contained specific marker genes from the other nucleus. A unilateral transfer of gene copies between nuclei thus seems to be associated with mound cell differentiation. The genetic change is not the result of mutation and also is distinguished from somatic recombination[6], but is similar to single factor transfer[7]. The transferred copies are stably integrated into the recipient genome and subsequently behave as mendelian genes.

S. commune grows as a dikaryon, cells with two compatible haploid nuclei, and genetic techniques have been well developed such that the two nuclei can be marked with a variety of gene mutations of known linkage relationships[8]. Eventually, cell differentiation occurs in various areas of the dikaryotic colony, resulting in the formation of multicellular, aerial reproductive structures. The two haploid nuclei can be induced to dissociate and form uninucleate cells when grown on a bile salt-containing medium[9]. Individual nuclei of dikaryons recovered in this way can be crossed with tester strains and their constellation of genetic markers determined. Genetic analysis involves collecting meiotically produced spores from the reproductive structures and carrying out random spore genetic analysis.

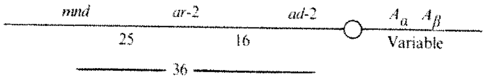


Fig. 1 *mnd* linkage group with loci relevant to the present study.

The evidence reported here involves the recessive mutant allele *mnd*, whose phenotype is observed as abnormal growths called 'mounds'[10]. These structures develop in some monokaryotic strains which possess the *mnd* allele and in all dikaryotic strains that are homoallelic for *mnd*[9]. Mounds are characteristically large aerial hemispherical bodies of compacted hyphae that exhibit properties of autonomous and indeterminate growth. At present it is not clear whether mounds represent abnormal reproductive structures (the mushroom-like fruiting bodies) or abnormal vegetative growth. We have previously presented evidence that mounds which develop on an initially heteroallelic (*mnd* + *mnd*$^+$), dikaryotic colony are composed of homoallelic (*mnd* + *mnd*) dikaryotic cells and that non-mound mycelium remaining in such colonies is composed of heteroallelic cells[9].

The behaviour of the *mnd* gene in dikaryons heteroallelic for *mnd* was examined in relation to other genes both linked and unlinked to *mnd*. The linkage relationship of the mutant alleles, *ad-2* and *ar-2*, to the *A* mating genes have been previously established[11]. The linkage relationships of *mnd* to *ad-2* and *ar-2* shown in Fig. 1 were generated by analysis of 130 random spore progeny from the cross: *A41 B41 mnd*$^+$ *ad-2 ar-2* × *A47 B42 mnd ad-2*$^+$ *ar-2*$^+$ (data not shown). Other mutant alleles used, *ni-3*, *ar-6* and *ur-1* as well as the *B* mating type locus, are located on separate chromosome arms[11] and are consequently unlinked to *mnd* and to each other (Table 1). The methods for scoring *mnd*, mating type genes and auxotrophic marker genes have been described previously[9].

A replicate series of 55 crosses of the type P1 (*A47 B42 mnd ad-2 ni-3 ur-1*$^+$ *ar-6*$^+$) × P2 (*A43 B41 mnd*$^+$ *ad-2*$^+$ *ni-3*$^+$ *ur-1 ar-6*) produced the expected dikaryotic colonies on which mounds developed. Mound tissue was subcultured from each of these colonies and grown on a medium containing bile salts to form numerous nuclear dissociations producing uninucleate cells. All the recovered monokaryons fell into two types (Table 1). One type was genetically identical to the original *mnd* parent strain (P1). The other was genetically similar to the *mnd*$^+$ parent strain (P2) with the exception that it had acquired *mnd* and *ad-2* alleles apparently replacing the corresponding alleles originally present. It thus seems that there is a donor nucleus (*mnd*) and a recipient nucleus (initially *mnd*$^+$) with respect to the genetic transfer that occurred. In addition, only genes linked to *mnd* showed translocation.

Table 1 Monokaryotic component strains recovered from mound and non-mound dikaryotic cells

Dikaryotic nuclei: P1 (*A47 B42 mnd ad-2 ni-3 ur-1*$^+$ *ar-6*$^+$) and *P2* (*A43 B41 mnd*$^+$ *ad-2*$^+$ *ni-3*$^+$ *ur-1 ar-6*) Recovered monokaryotic cells*			
From mound tissue		From non-mound tissue	
Nuclear types:	No. recovered	Nuclear types:	No. recovered
A47 B42 mnd ad-2 ni-3 ur-1$^+$ *ar-6*$^+$	418	*A47 B42 mnd ad-2 ni-3 ur-1*$^+$ *ar-6*$^+$	361
*A43 B41 mnd** *ad-2** *ni-3*$^+$ *ur-1 ar-6*	132	*A43 B41 mnd*$^+$ *ad-2*$^+$ *ni-3*$^+$ *ur-1 ar-6*	129

Genetic marker abbreviations used above and in Fig. 1. *mnd*, Mound allele; *ni-3*, nicotinamide requiring; *ar-6*, arginine requiring; *ad-2*, adenine requiring; *ur-1*, uracil requiring. *A* factor, the *A* incompatibility factor which regulates aspects of the dikaryotic phenotype; *B* factor, the *B* incompatibility factor which regulates aspects of the dikaryotic phenotype.

* Recovered strains designated *mnd** were originally the *mnd*$^+$ component of the parent dikaryon, but are now *mnd*. Similarly for *ad-2**.

The frequency with which genes linked to *mnd* co-transferred with *mnd* in dikaryotic mound tissue was examined in subcultures from 125 replicate crosses of the type *A51 B51 mnd ad-2 ar-2* × *A41 B41 mnd*$^+$ *ad-2*$^+$ *ar-2*$^+$. Each of the 125 mound isolates was scored for homozygosity of genetic markers as observed by growth on appropriately substituted minimal media. As expected, each of the mound isolates had become homozygous for the *mnd* allele and 115 of the isolates had also become homozygous for the *ad-2* and *ar-2* alleles. The other 10 mound isolates remained heterozygous for these alleles. The mating type loci also remained heterozygous in all the isolates, as indicated by the continued dikaryotic state of the nuclei within the cells. Thus, the co-transfer of genes linked to *mnd* is frequent but variable: in 9% of the above events the markers linked to *mnd* were not observed to co-transfer with *mnd*.

We have also observed in other crosses that 'co-transferred' genes can be either wild type or mutant, that is, *ar-2*$^+$ and *ad-2*$^+$ have also replaced their allelic counterparts when co-transferred with *mnd* (unpublished). Such evidence makes it unlikely that it is a *trans*-acting 'signal' that is transferred, that is, a signal from *mnd* which converts *mnd*$^+$ to *mnd*, by something akin to chromosomal 'inactivation'.

Although a positive selection for the transfer of genes in phenotypically heteroallelic non-mound vegetative cells has not been developed, the sampling and testing of thousands of these cells have not revealed the transfer of *mnd* or any other marked genes[9,12]. The development of a selection system using recessive drug-resistance markers is in progress to discern whether there is any transfer of genes from *mnd*$^+$ nuclei or if this is strictly a *mnd*-dependent unidirectional transfer.

We conclude that in the binucleate heterokaryon phase of *S. commune* there is a process, as yet uncharacterized, by which copies of genes from one nucleus can be unilaterally transferred to the other nucleus of the dikaryon. The transferred genes seem to substitute for their corresponding alleles in the recipient nucleus, and subsequently have been demonstrated to behave in a mendelian fashion[9,12]. The previous alleles in the recipient nucleus are apparently lost. The donor nucleus is not detectably altered during the transfer process. All these observations suggest that gene transfer involving the *mnd* region of the *A* mating factor chromosome occurs without nuclear fusion.

This work was supported by the College of Agricultural and Life Sciences, University of Wisconsin, Madison, and by NSF grant PCM77-05581 to T.J.L.

Received 2 January; accepted 16 March 1981.

1. Bukhari, A. I., Shapiro, J. A. & Adhya, S. L. (eds) *DNA Insertion Elements. Plasmids and Episomes* (Cold Spring Harbor Laboratory, New York, 1977).
2. Kearney, J., Paige, C. & Tonegawa, S. *Science* **209**, 1366–1369 (1980).
3. Simon, M., Zieg, J., Silverman, M., Mandel, G. & Doolittle, R. *Science* **209**, 1370–1374 (1980).
4. Zambryski, P. *et al. Science* **209**, 1385–1391 (1980).
5. Hicks, J. B., Strathern, J. N. & Herskowitz, I. in *DNA Insertion Elements, Plasmids and Episomes* (eds Bukhari, A. I., Shapiro, J. A. & Adhya, S. L.) 457–462 (Cold Spring Harbor Laboratory, New York, 1977).
6. Pontecorvo, G. *A. Rev. Microbiol.* **10**, 393–400 (1956).
7. Ellingboe, A. H. & Raper, J. R. *Genetics* **47**, 85–98 (1962).
8. Raper, J. R. *Genetics of Sexuality in Higher Fungi* (Ronald Press, New York, 1966).
9. Leonard, T. J., Gaber, R. & Dick, S. *Genetics* **89**, 685–693 (1978).
10. Leonard, T. J. *Proc. natn. Acad. Sci. U.S.A.* **72**, 4626—4630 (1975).
11. Raper, J. R. & Hoffmann, R. M. in *Handbook of Genetics* Vol. 1, 597–626 (1974).
12. Leonard, T. J., Dick, S. & Gaber, R. *Genetics* **88**, 13–26 (1978).

A small RNA complementary to an intervening sequence is produced late in SV40 infection

Rodolfo Campos, Susan Jovanovich & Luis P. Villarreal*

Department of Microbiology and Immunology, University of Colorado Health Sciences Center, 4200 East Ninth Avenue, Denver, Colorado 80262, USA

The post-transcriptional modifications that affect simian virus 40 (SV40) mRNA are similar to those undergone by most eukaryotic mRNAs and include 5′-end capping, 3′ polyadenylation and splicing of mRNA precursors[1]. Late in SV40 infection unspliced 19S RNAs are found predominantly in the nucleus whereas the spliced 19S RNAs are mainly cytoplasmic[2,3]. Nuclear and cytoplasmic 19S RNAs differ in that all the cytoplasmic species have lost a 32-nucleotide intervening sequence located between 0.760 and 0.765 map units[3,4]. Previous transcriptional studies of SV40 deletion mutants implied that splicing of the late nuclear transcripts and transport of RNA into the cytoplasm were coupled events[5]. More recent studies, however, demonstrate that transport of unspliced 19S RNA can occur when the sequences encompassing the 32-nucleotide intervening sequence (0.760–0.765 map units) are deleted[6]. We have previously proposed that this small 19S RNA intervening sequence may be involved in the retention of the unspliced RNA within the nucleus[3]. One hypothesis to account for recognition of the 32-nucleotide intervening sequence by a nuclear retention system postulates the existence of a small nuclear RNA complementary to the intervening sequence. We describe here the initial characterization of a small nuclear RNA made late in SV40 infection which is complementary to the DNA of the 19S intervening sequence. Several host cytoplasmic RNAs also hybridize to this intervening sequence.

The initial experimental step was the construction of a hybrid plasmid DNA carrying the 32-nucleotide intervening sequence (IVS) DNA from SV40 DNA (0.760–0.765 map units), and removal of the 32-base pair fragment after appropriate restriction enzyme cleavage to the pBR325-DNA-plasmid vector (Fig. 1). The resulting product was called PBR325/IVS and was attached to diazobenzyloxymethyl (DBM)-paper according to the procedure described by Alwine *et al.*[7]. RNA present in the nucleus of SV40-infected cells during the late infection phase was extracted by an isotonic method[3,8]. This nuclear RNA was labelled at the 5′ end[9] and hybridized against the PBR325/IVS DNA attached to the DBM-paper. Hybridized RNA was eluted and analysed by electrophoresis in a 12% acrylamide gel in denaturing conditions (8 M urea). The results (Fig. 2) show the presence of a small RNA of ~60 nucleotides which is present in the nucleus of infected cells but not in that of uninfected cells (Fig. 2, lane *a*). In a control experiment using vector DNA without IVS DNA attached to DBM-paper, no RNA was retained by the paper (data not shown). The sensitivity of this small RNA to digestion with RNase is shown in Fig. 3, lanes *a*, *b*. These results indicate that the small RNA found in the nucleus of SV40-infected cells is in fact complementary to the IVS DNA sequence. In addition to the 60-nucleotide RNA (henceforth referred to as ivsRNA), the nuclear fraction contains a very minor RNA band corresponding to ~38 nucleotides.

The existence of small RNA molecules within the nucleus (snRNA) of eukaryotic cells is well established and many of these snRNAs have been characterized[10,11]. However, the induction of a small nuclear RNA complementary to an intervening sequence has not been previously described. The function of the snRNAs is not established, although it has been proposed that they are involved in the mRNA splicing process.

The small virus-induced nuclear RNA described here possesses two marked differences from previously described snRNAs: it is generally smaller, 60 compared with ~200 nucleotides, and it does not have a 5′-terminal cap, as this ivsRNA can be labelled at the 5′ end. Although post-transcriptional cleavage should be considered, this suggests that the synthesis of the SV40 sRNA differs from that of most eukaryotic snRNAs.

Our hybridization selection procedure for the small RNAs did not use sequences adjacent to the 32-nucleotide intervening sequence, and the conditions of hybridization are reasonably stringent. Therefore, due to the short length of the hybrids, all the small RNAs we described must have almost complete homology to the intervening sequence. Although our search for these small RNAs stems from the prediction of a recognition system for the retention of nuclear transcripts, their existence does not necessarily confirm this hypothesis and other explanations are possible[12–14].

The subcellular distribution of small RNAs complementary to the intervening sequence is also shown in Fig. 2. The 60-nucleotide sRNA is clearly more abundant in the nuclear fraction. The paranuclear fraction also contains the 60-nucleotide

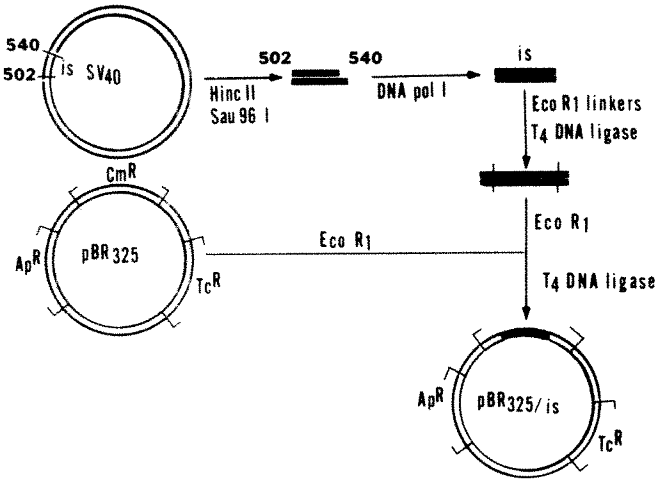


Fig. 1 Scheme for the construction of pBR325/IVS: SV40 DNA was cut with *Hinc*II and *Sau*96I restriction enzymes, and the small 38-base pair fragment generated by this cleavage was gel purified. After electroelution from the gel, the fragment was treated with DNA polymerase I to fill the *Sau*96I terminus as described elsewhere[18]. *Eco*RI linkers were joined to the 38-base pair fragment with T_4 ligase, generating a 50-base pair fragment[19], and the pBR325-DNA and the 50-base pair fragment DNA cut with *Eco*RI. The 5′-cohesive termini were joined to each other with T_4 ligase as previously described[20]. The plasmid DNAs were used to transform *Escherichia coli* HB101 cells[21], and the transformants selected for growth in the presence of ampicillin. Transformed cells carrying pBR325/IVS DNA were then selected for loss of chloramphenicol resistance. The structure of the plasmic DNA in candidate colonies was confirmed by further restriction analysis and subsequently by DNA sequencing according to Maxam and Gilbert[9].

0028-0836/81/220344—03$01.00

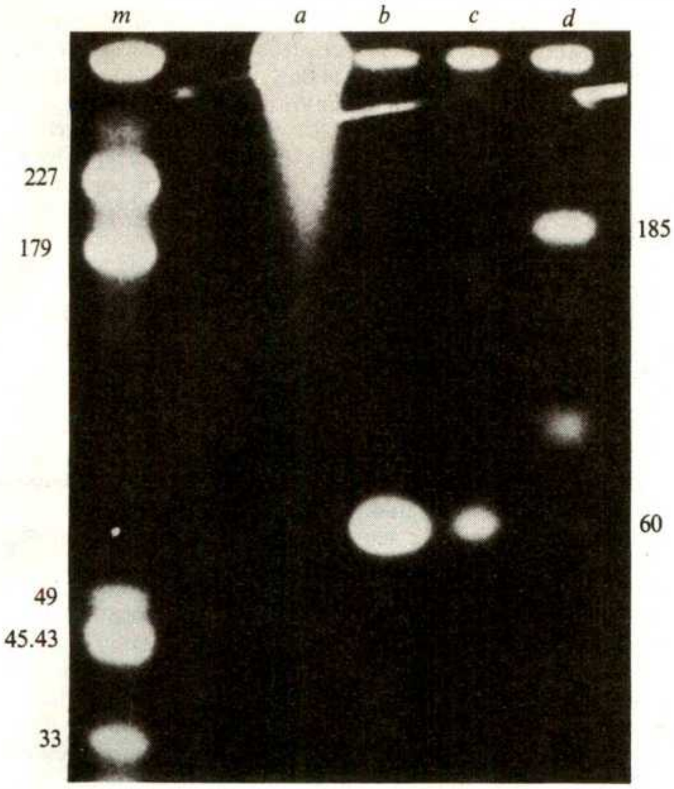


Fig. 2 Gel analysis of ivsRNA. Semiconfluent monolayers were infected at a high multiplicity of infection (MOI = 10), and 48 h later RNA was extracted using isotonic procedure as described previously[3,8]. Briefly, the cell monolayers, after washing with TSM (30 mM Tris-HCl *p*H 7.6, 150 mM NaCl, 1.5 mM $MgCl_2$) and lysis with TSM plus 0.2% NP40 on ice, were scraped off the plates and centrifuged (1,000*g*, 3 min). The cytoplasmic supernatant was decanted and the residual nuclei removed by a second centrifugation. The 1,000*g* pellets (nuclear and paranuclear fraction) were resuspended with TSM–NP40, pelleted again and resuspended with TSM–NP40, 0.2%, plus 0.05% deoxycholate and centrifuged again at 1,000*g*. This supernatant (paranuclear fraction) was decanted and the nuclear pellet resuspended in TSM–NP40. All fractions were mixed with extraction buffer (10 mM Tris-HCl *p*H 7.8, 1 mM EDTA, 350 mM NaCl, 2% SDS and 7 M urea), extracted twice with H_2O and saturated 1:1 phenol/chloroform and ethanol precipitated. Extracted RNAs were digested with DNase[3], dephosphorylated, 5′-end labelled using $[\gamma\text{-}^{32}P]$ATP[9] then ethanol precipitated. Dry pellets from each fraction were resuspended in hybridization buffer (50% formamide, 100 mM PIPES, *p*H 6.8, 600 mM NaCl, 1 mM EDTA, 0.1% SDS) and hybridized against pBR325/IVS DNA fixed to DBM-paper preformed as described elsewhere[7]. After 48 h at 45 °C the hybrids were washed at 45 °C three times with 40% formamide, 4 × SSC, 0.1% SDS and three times more with 2 × SSC at room temperature. RNA elution from the pBR325/IVS DNA-paper was done twice at 65 °C for 15 min with 100% formamide and the resultant eluted material was precipitated with ethanol. The pellets were resuspended in TE (10 mM Tris-HCl *p*H 7.6, 1 mM EDTA) and equal volumes of formamide–dye mixture were added. After 2–3 min at 100 °C, the samples were loaded on a 12% acrylamide gel with 8 M urea in TBE (50 mM Tris-borate *p*H 8.3, 1 mM EDTA) and run for 14–16 h at 85 V. Wet gels were exposed with Kodak XR-5 film at −70 °C for 24 h. *a*, Nuclear RNA from uninfected cells; *b*, nuclear RNA from infected cells; *c*, paranuclear RNA from infected cells; *d*, cytoplasmic RNA from infected cells; *m*, SV40 *Hae*III markers.

species but at substantially reduced levels. The small RNA species present in the cytoplasmic fraction differ significantly from those in the nucleus; the most abundant is 185 nucleotides long and is present in uninfected cells. Several minor species (see Fig. 3) of ~160, 87, 77 and 73 nucleotides length are also observed. The major cytoplasmic species is much longer and more abundant than the nuclear species (185 compared with 60 nucleotides). Because the cytoplasmic ivsRNAs are all longer than the nuclear species, and are mostly present in uninfected cells, leakage of nuclear species cannot account for any of the cytoplasmic ivsRNAs seen. The function of these RNAs is unknown but their presence in the cytoplasm may imply an involvement with the translational machinery. Small cytoplasmic RNAs have previously been observed during adenovirus[15] and SV40 infection[16]. A 65-nucleotide cytoplasmic RNA that arises late in SV40 infection and is complementary to early mRNA at 0.21 map units (SAS RNA) has been described by Alwine and Khoury[17], but unlike the nuclear ivsRNA described here, their SAS RNA has no apparent association with splice junctions or intervening sequences. Furthermore, the existence of a small RNA which will hybridize to SV40 DNA near or at 0.76 map units has been independently observed by others (J. Alwine, personal communication). Clearly, more information on the sequence and function of these SV40 ivsRNAs is necessary to establish their role during infection.

We thank Susan Carr for technical assistance. R. C. is a fellow from the Consejo de Investigaciones Cientificas Y Tecnicas, Republica Argentina. This work was supported by PHS grant CA-25819 from the NCI to L.P.V.

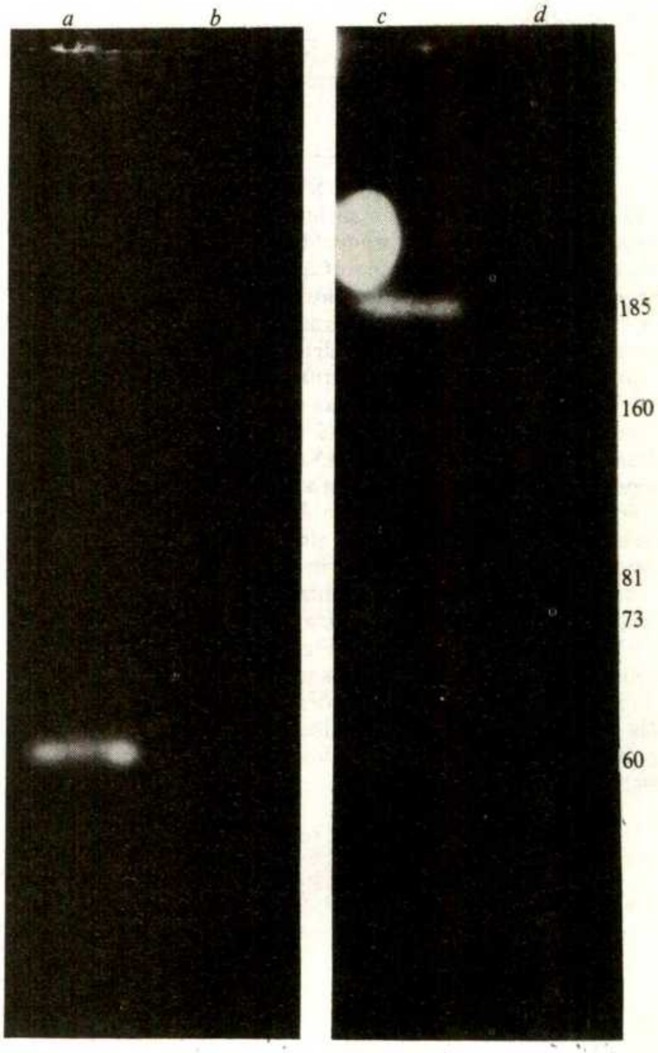


Fig. 3 RNase sensitivity of small RNAs. RNAs from nuclear and cytoplasmic fractions of infected cells were extracted and selected as described in Fig. 2 legend. Nuclear and cytoplasmic small RNAs were then treated with RNase (RNase A, 0.1 mg ml^{-1}, RNase T, 50 U ml^{-1}, 37 °C, 45 min) in 10 mM $MgCl_2$, 10 mM Tris-HCl *p*H 7.6, 2.5 mM dithiothreitol and 0.25 mg ml^{-1} gelatin. *a*, Nuclear RNA; *b*, nuclear RNA plus RNase; *c*, cytoplasmic RNA; *d*, cytoplasmic RNA plus RNase.

Received 3 February; accepted 11 March 1981.

1. Ziff, E. *Nature* **287,** 491–499 (1980).
2. Lai, D. R., Dhar, R. & Khoury, G. *Cell* **14,** 971–982 (1978).
3. Villarreal, L. P. & Jovanovich, S. *J. Virol.* **36,** 595–600 (1980).
4. Ghosh, P. *et al. J. molec. Biol.* **126,** 813–846 (1978).
5. Gruss, P. *et al. Proc. natn. Acad. Sci. U.S.A.* **76,** 4317–4321 (1979).
6. White, R. T., Berg, P. & Villarreal, L. P. *J. Virol.* (submitted).
7. Alwine, J. C., Kemp, D. J. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5350–5354 (1977).
8. Goldenberg, C. J. & Raskas, H. J. *Cell* **16,** 131–138 (1979).
9. Maxam, A. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74,** 560–564 (1977).
10. Weinberg, R. & Penman, S. *J. molec. Biol.* **38,** 289–304 (1968).
11. Shibata, H. *et al. J. biol. Chem.* **250,** 3909–3920 (1975).
12. Jelinek, W. & Leinwand, L. *Cell* **15,** 205–214 (1978).
13. Lerner, M. *et al. Nature* **283,** 220–224 (1980).
14. Roberts, R. *Nature* **283,** 132–133 (1980).
15. Mathews, M. B. & Pettersson, V. J. *J. molec. Biol.* **119,** 293–328 (1978).
16. Alwine, J. C., Dhar, R. & Khoury, G. *Proc. natn. Acad. Sci. U.S.A.* **77,** 1374–1383 (1980).
17. Alwine, J. C. & Khoury, G. *J. Virol.* **36,** 701–708 (1980).
18. Seeburg, P. *et al. Nature* **270,** 486–494 (1977).
19. Heyneker, H. L. *et al. Nature* **263,** 748–752 (1976).
20. Dugaiczyk, A., Boyer, H. & Goodman, H. *J. molec. Biol.* **96,** 171–184 (1975).
21. Wensink, P. *et al. Cell* **3,** 315–325 (1974).

Identification of the SV40 agnogene product: a DNA binding protein

Gilbert Jay*, Shigeko Nomura*, Carl W. Anderson† & George Khoury*

* Laboratory of Molecular Virology, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20205, USA
† Biology Department, Brookhaven National Laboratory, Upton, New York 11724, USA

The proximal late region of the SV40 genome (map positions 0.72–0.76) contains an open nucleotide reading frame[1,2] called the agnogene[3] because its product had so far not been detected. We report here the detection of a novel protein of molecular weight (M_r) 7,900 (61 amino acids) in cells lytically infected by SV40; its identity has been confirmed by comparing the amino-terminal sequence with that predicted from the SV40 nucleotide sequence. This so-called agnoprotein accumulates late in the lytic cycle and has a relatively short half life of ~2 h. The highly basic nature of the protein and its affinity for both single-stranded and double-stranded DNA suggests that it may have a regulatory role in nucleic acid–protein interactions.

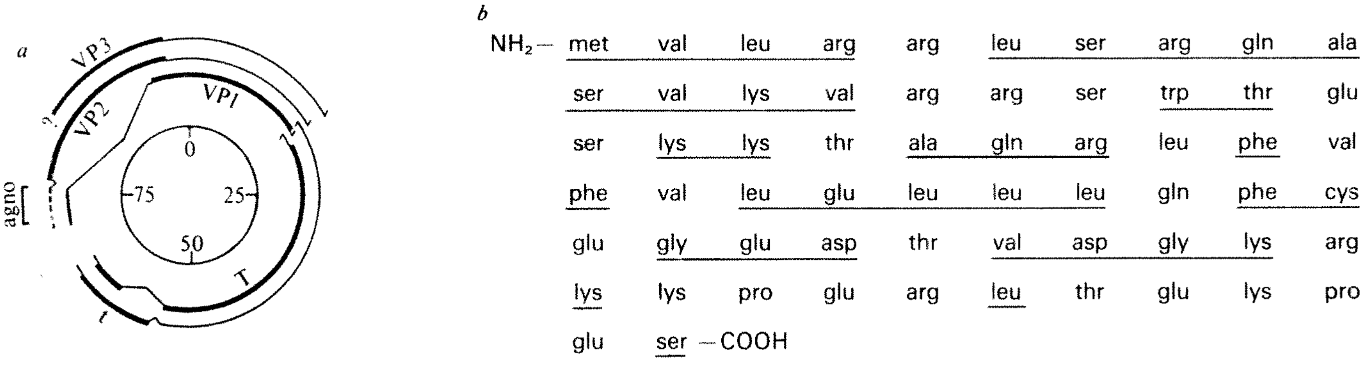
For the SV40 genome, up to 20% of the DNA sequence encodes no known protein[1,2]. A significant portion of this non-coding sequence seems to exist in regulatory structures but at least one segment in the proximal late region of the genome contains an open reading frame (agnogene) capable of encoding a 62-amino acid polypeptide. Figure 1 shows the agnogene, located between positions 0.72 and 0.76 on the SV40 genome (*a*) and the amino acid sequence of its product (*b*). The sequence has a large number of the amino acids arginine, lysine and leucine; the only methionine codon present is that of the initiator AUG.

To detect this protein product, SV40-infected African green monkey kidney cell cultures were labelled with ^{14}C-arginine, and proteins were extracted late in the lytic cycle. Figure 2 shows a protein band migrating in the position expected for the agnoprotein (7,900 M_r) in wild-type SV40-infected cells, which is absent from both mock-infected cells and cells infected by a viable deletion mutant of SV40 (dl 2304) which contains a deletion over the entire region of the agnogene (unpublished data). A protein migrating in the same position was detected when cells were labelled with ^{14}C-leucine, but not with ^{35}S-methionine (data not shown).

Two types of experiment suggested that the 7,900 M_r (7.9K) polypeptide observed in Fig. 2 was in fact the agnoprotein: (1) cell-free translation of late lytic SV40 RNA selected by hybridization to SV40 DNA-containing filters gave a polypeptide which co-migrated with the species identified in Fig. 2 (data not shown). (2) A two-nucleotide insertion mutant of SV40 which shifts the reading frame for the putative agnoprotein produced an appropriately truncated polypeptide (data not shown). Amino-terminal peptide sequencing (Fig. 3) provided direct evidence for the identity of the agnogene product. Extracts from lytically infected monkey kidney cells, labelled with either ^{14}C-leucine or ^{14}C-arginine, were prepared by affinity chromatography, followed by SDS-polyacrylamide gel electrophoresis; the 7.9K band was eluted and subjected to N-terminal sequence analysis[4]. An amino acid sequence through the first 16 Edman cycles confirms the identity of the agnogene product. The observed sequence, xxx-Leu-Arg-Arg-Leu-xxx-Arg, is present only once in the entire viral genome, at the beginning of the agnogene. An almost quantitative recovery of radiolabelled counts suggests that little if any of the polypeptide contains a blocked amino terminus. Furthermore, the positions of the confirmed amino acids compared with the putative sequence of the polypeptide (Fig. 1) indicate that the initiating methionine is quantitatively removed, leaving valine as the N-terminal amino acid.

To investigate the time course of appearance of the 7.9K polypeptide, monkey kidney cell cultures were labelled at various times after infection (see Fig. 4); the polypeptide appeared only late after lytic infection and thus seems to parallel the appearance of the virion polypeptides (Fig. 4*a*). For comparison, the SV40 tumour antigens were immunoprecipitated from the same set of samples using a broad spectrum anti-tumour serum (Fig. 4*b*). Although the large T and small t antigens seem to be synthesized in greater amounts late in infection, their presence is also easily detected at early stages. As it was conceivable that some agnoprotein was made early in the lytic cycle at levels below our detection, we tried to find an antiserum to the agnoprotein to increase the efficiency of its detection. None of the available antisera, including different anti-tumour sera and sera produced against SDS-disrupted virions, were able to immunoprecipitate the agnoprotein.

Data from pulse-chase experiments performed 48 h post-infection (not shown) indicate that the 7.9K polypeptide decays



NH$_2$— met val leu arg arg leu ser arg gln ala
ser val lys val arg arg ser trp thr glu
ser lys lys thr ala gln arg leu phe val
phe val leu glu leu leu leu gln phe cys
glu gly glu asp thr val asp gly lys arg
lys lys pro glu arg leu thr glu lys pro
glu ser —COOH

Fig. 1 Coding region and amino acid sequence of the putative product of the SV40 agnogene. *a*, Map location of the agnogene relative to the known SV40 mRNAs. The heavy lines indicate the coding sequences within each of the mRNAs. *b*, Amino acid sequence of the agnogene product as deduced from the known nucleotide sequence. Amino acid residues which are shared with the BKV agnogene product are underlined.

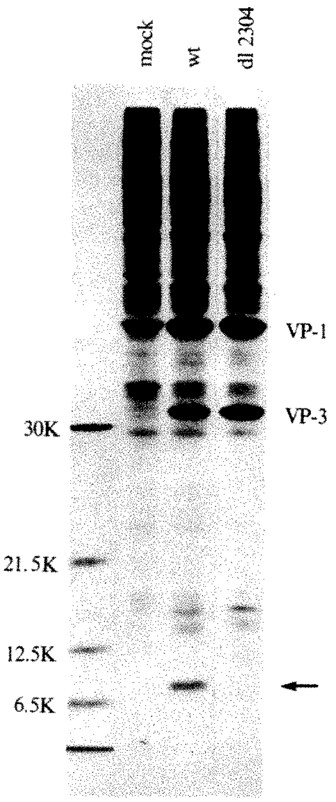


Fig. 2 Identification of the SV40 agnogene product. Secondary cultures of African green monkey kidney (AGMK) cells were infected with either wild-type (wt) SV40 or a deletion mutant dl 2304, at a multiplicity of 10–20 plaque-forming units (PFU) per cell. At 48 h post-infection, the cultures were radiolabelled with ^{14}C-arginine (350 mCi mmol^{-1}) at 16.7 μCi ml^{-1} for 2 h. Cells were collected and lysed in buffer A (50 mM Tris-HCl *p*H 7.5, 150 mM NaCl, 1% Triton X-100, 1% sodium deoxycholate, 0.1% SDS, 25 μg ml^{-1}, L-1-tosylamide-2-phenylethylchloromethyl ketone). After centrifugation at 100,000 *g* for 30 min, the supernatant was analysed on a 12.5% SDS-polyacrylamide gel as described elsewhere[20]. The gel was fluorographed and dried, and an autoradiogram obtained by exposing the dried gel to Kodak XR-2 films at −80 °C. The molecular weight markers used were carbonic anhydrase (30 K), soybean trypsin inhibitor (21.5 K), cytochrome *c* (12.5 K) and aprotinin (6.5 K). The arrow indicates the position of the agnoprotein.

with a half life of $\sim$2 h which contrasts with that of T antigen ($\sim$10 h) and VP-1 (15 h) determined in parallel. This short half life, together with the relative ease of labelling of the agnoprotein, suggest a rapid synthetic rate and indicate a potential regulatory function rather than a structural role for the agnoprotein.

The highly basic nature of the agnoprotein (arginine + lysine = 25%) suggested that it may bind to DNA and/or RNA; thus a binding assay was done using single-stranded (ss) and double-stranded (ds) calf thymus DNA with lytically infected cell extracts which contained both the agnoprotein and the known DNA binding protein, SV40 T antigen[5–8]. Figure 5*a* shows that almost all the input agnoprotein was retained on either ss or dsDNA columns at low salt concentration (0.1 M NaCl) and could be quantitatively eluted with high salt (1.0 M NaCl). In contrast (Fig. 5*b*) $\sim$50% of the SV40 T antigen did not bind to either the ss or dsDNA in identical conditions (0.1 M NaCl). None of the SV40 t antigen (previously shown not to be a DNA binding protein[9]) was retained in these conditions. Thus, unlike SV40 T antigen, which seems to be composed of heterogeneous populations with respect to DNA binding properties, the agnogene protein homogeneously and quantitatively binds both ss and dsDNA. The agnoprotein, and the bound fraction of T antigen, elute at 0.3 M NaCl (data not shown).

These data indicate that the SV40 agnoprotein is synthesized in large quantities late in the lytic cycle and that has a remarkably short half life. The observation that mutants with deletions within the agnogene region are viable[10–14] suggests that, at least in a tissue culture system, the agnoprotein, like the t antigen encoded by the early region of SV40, is not essential. However, several studies have shown that mutants with deletions in this region often do not grow as well as wild-type virus[15]. A small genome like SV40, subject to rapid evolutionary pressures, would not contain sequences which were of little or no advantage. Thus we can speculate that the agnoprotein may act simply to enhance the efficiency of SV40-specific functions. Alternatively, its role in tissue culture systems may be substituted by a host cell protein (albeit with a lower efficiency). A significant role for the SV40 agnoprotein is suggested by the conservation

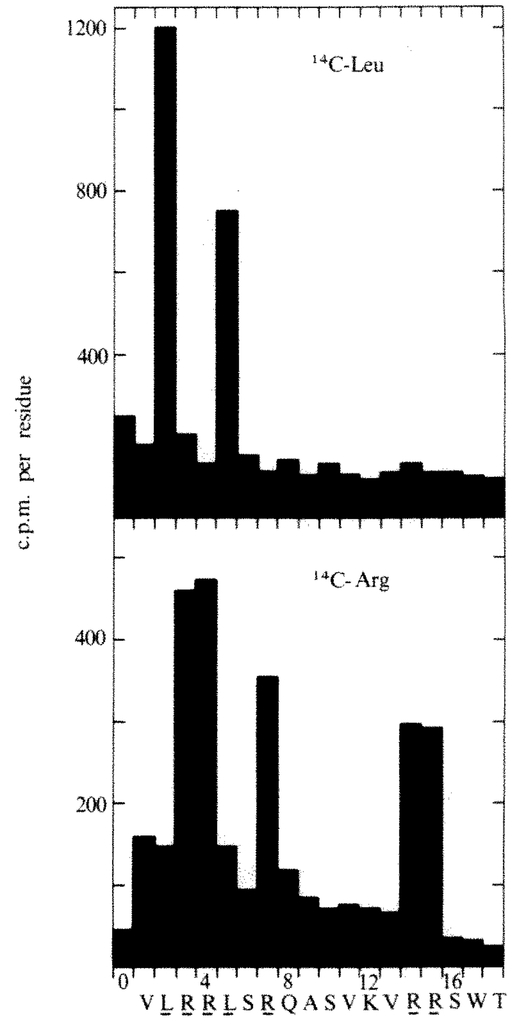


Fig. 3 Partial amino-terminal sequence determination of the SV40 agnogene product. Lysates of SV40-infected cells radiolabelled with ^{14}C-arginine or ^{14}C-leucine were enriched for the agnoprotein by affinity chromatography on a calf thymus DNA–cellulose column (see Fig. 5 legend), followed by electrophoresis on a 12.5% SDS-polyacrylamide gel. The position of the agnoprotein was identified by autoradiography of the dried gel and this region excised and eluted electrophoretically. The eluted samples were applied to a Beckman 890C automated protein sequencer together with 2 mg of apomyoglobin carrier protein and 3 mg of polybrene. About 6,500 c.p.m. of ^{14}C-arginine-labelled product and 7,400 c.p.m. of ^{14}C-leucine-labelled product were applied to the sequencer. The amount of radioactivity released after each Edman cycle is shown. The expected amino acid sequence of the SV40 agnoprotein amino-terminus is given in the single letter code at the bottom of the figure; identified residues are underlined. The initial yield, estimated from the apparent repetitive yield and the expected amino acid composition, was 76% for the arginine-labelled sample and slightly greater than 100% for the leucine-labelled sample.

of this sequence within the genome of the human papovavirus BKV[16,17], an evolutionary variant of SV40. About two-thirds of the amino acids in the putative BKV agnogene are identical to those of the SV40 protein (see Fig. 1*b*).

The location of the coding sequence for the SV40 agnogene within the leader of the major late mRNAs may suggest a role in transcription or processing of the late SV40 mRNA. In this regard, recent studies have indicated that an attenuator of late SV40 mRNAs is situated in the proximal late region[18]. It is possible that the agnoprotein is translated from an attenuated RNA and/or affects this process of attenuation.

Given the affinity of the agnoprotein for DNA and its basic nature, it is possible that, like histones, it might be associated with the viral nucleic acid core. Although preliminary experiments suggest that there is no detectable agnogene product in the SV40 virion (data not shown), this does not exclude the possibility that the agnoprotein has a transient function in virion assembly.

We do not know which RNA species encode(s) the agnoprotein. Of the two major late SV40 polyadenylated transcripts (16S and 19S RNAs), the 16S mRNA is the more likely candidate because of its abundance and the fact that the agnogene coding sequence is found in its leader (5′ exon). A considerable body of data has led to the supposition that 5′-terminal AUG codons are preferentially selected by eukaryotic translational

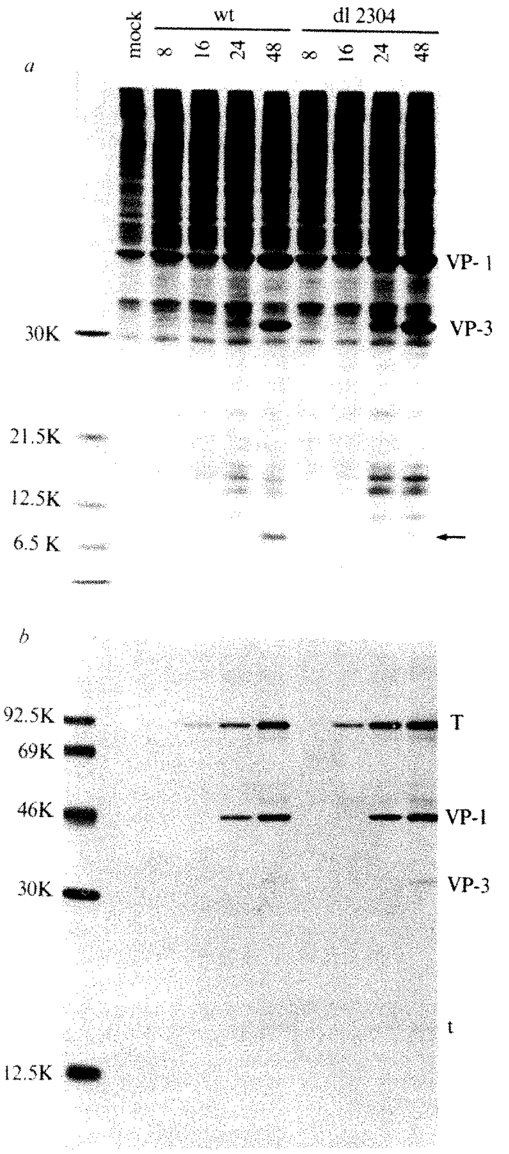


Fig. 4 Time course of the synthesis of the SV40 agnogene product. Cultures of AGMK cells were infected with either wt SV40 or dl 2304. At various times post-infection, cells were radiolabelled for 1 h with ^{14}C-arginine and collected as described in Fig. 2 legend. Aliquots of the cell extracts were analysed either directly (*a*), or after immunoprecipitation with an SV40 anti-tumour serum (*b*) by electrophoresis on a 12.5% SDS-polyacrylamide gel, followed by fluorography and exposure for 12 h. The molecular weight markers used in *a* were the same as those described in Fig. 2 legend; those in *b* were phosphorylase *b* (92.5 K), bovine serum albumin (69 K), ovalbumin (46 K), carbonic anhydrase (30 K) and cytochrome *c* (12.5 K). The arrow indicates the position of the agnoprotein.

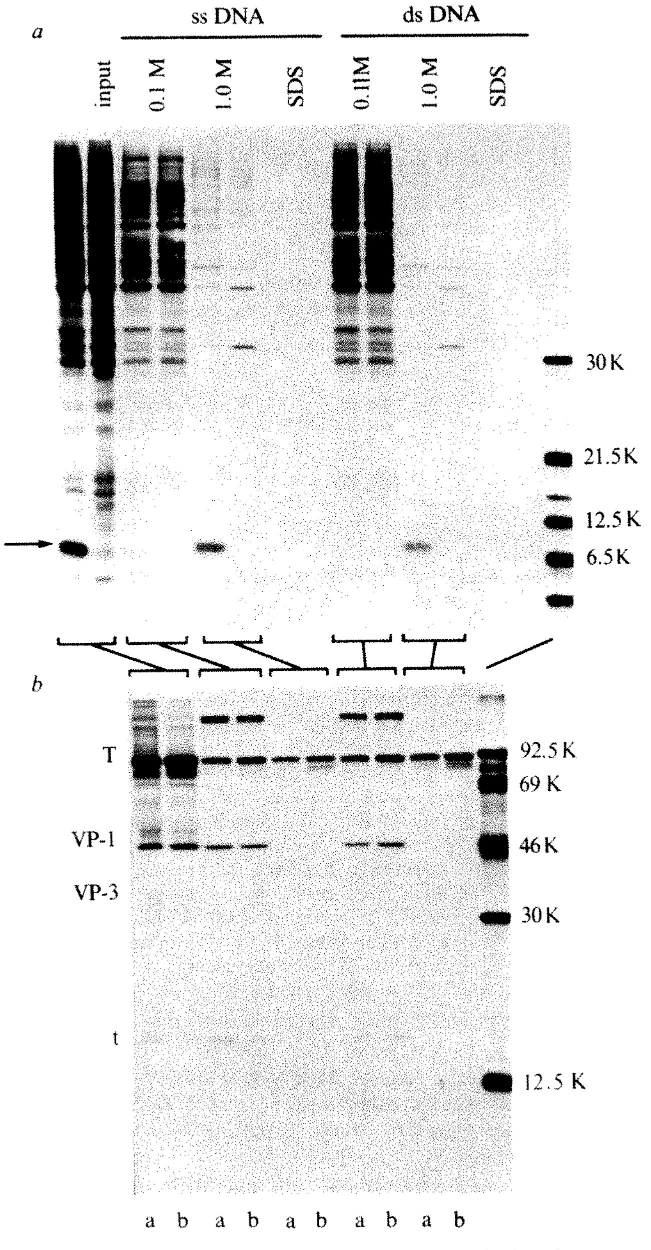


Fig. 5 Binding of the SV40 agnogene product to DNA. Cultures of AGMK cells, infected with either wt SV40 (a) or dl 2304 (b) at 10–20 PFU per cell, were labelled for 2 h with ^{14}C-arginine at 48 h post-infection. Cells were collected and lysed in buffer B (Tris-buffered saline *p*H 8.0 containing 1% Triton X-100 and 25 μg ml^{-1} L-1-tosylamide-2-phenylethylchloromethylketone). The extracts were adjusted to *p*H 6.0 and loaded on to either denatured (ss) or native (ds) calf thymus DNA–cellulose columns, previously equilibrated with 5 mM sodium phosphate, *p*H 6.0/0.1 M NaCl/0.5% Nonidet P-40/10% glycerol. The columns were washed with the same buffer and the bound fractions were stepwise eluted, first with the same buffer containing 1 M NaCl then with buffer containing 2% SDS and 5% 2-mercaptoethanol. The various fractions were analysed either directly (*a*) or after immunoprecipitation with anti-tumour serum (*b*) in 12.5% SDS-polyacrylamide gels. The molecular weight markers were the same as those described in Fig. 4 legend. The arrow indicates the position of the agnoprotein.

systems (see ref. 19). But until now, the late 16S RNA had been considered to be an exception as the agnogene AUG precedes the AUG initiator codon which was known to be used for the translation of VP1. If the 16S mRNA is the translational template for the agnoprotein, the 5'-terminal AUG is a translational initiator. However, the late 16S SV40 RNA would then represent a unique polycistronic eukaryotic mRNA which can encode both the agnoprotein and VP1.

The absence of methionine residues from the agnoprotein as well as its rapid turnover rate may explain why it has not previously been detected. We have recently learned that other investigators have found a similar polypeptide in SV40-infected cells (P. Southern and P. Berg, J. Mertz, R. Margolsky and D. Nathans, personal communications).

This work was supported in part by the US Department of Energy. We thank Carol Prives for valuable suggestions.

Received 15 December 1980; accepted 5 March 1981.

1. Reddy, V. B. *et al. Science* **200,** 494–502 (1978).
2. Fiers, W. *et al. Nature* **273,** 113–120 (1978).
3. Dhar, R., Subramanian, K. N., Pan, J. & Weissman, S. M. *Proc. natn. Acad. Sci. U.S.A.* **74,** 827–831 (1977).
4. Anderson, C. W. & Lewis, J. B. *Virology* **104,** 27–41 (1980).
5. Carroll, R. B., Hager, L. & Dulbecco, R. *Proc. natn. Acad. Sci. U.S.A.* **71,** 3754–3757 (1974).
6. Reed, S. I., Ferguson, J., Davis, R. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **72,** 1605–1609 (1976).
7. Jessel, D. *et al. Cell* **8,** 535–545 (1976).
8. Tjian, R. *Cell* **13,** 165–179 (1978).
9. Prives, C. & Beck, Y. *INSERM Symp.* **69,** 175–188 (1977).
10. Mertz, J. E. & Berg, P. *Proc. natn. Acad. Sci. U.S.A.* **71,** 4879–4883 (1974).
11. Shenk, T. E., Carbon, J. & Berg, P. *J. Virol.* **18,** 664–671 (1976).
12. Subramanian, K. N. *Proc. natn. Acad. Sci. U.S.A.* **72,** 988–993 (1979).
13. Haegeman, G., van Heuverswyn, H., Gheysen, D. & Fiers, W. *J. Virol.* **31,** 484–493 (1979).
14. Villarreal, L., White, R. & Berg, P. *J. Virol.* **29,** 209–219 (1979).
15. Barkan, A. & Mertz, J. *J. Virol.* (in the press).
16. Seif, I., Khoury, G. & Dhar, R. *Cell* **18,** 963–977 (1979).
17. Yang, R. C. A. & Wu, R. *Science* **206,** 456–462 (1979).
18. Laub, O., Jakobovits, E. B. & Aloni, Y. *Proc. natn. Acad. Sci. U.S.A.* **77,** 3297–3301 (1980).
19. Kozak, M. *Cell* **22,** 7–8 (1980).
20. Jay, G., Jay, F. T., Chang, C., Friedman, R. M. & Levine, A. S. *Proc. natn. Acad. Sci. U.S.A.* **75,** 3055–3059 (1978).

Mutagenesis *in vitro* by depurination of ΦX174 DNA

Thomas A. Kunkel, Clyde W. Shearman & Lawrence A. Loeb

Joseph Gottstein Memorial Cancer Research Laboratory,
Department of Pathology, SM-30, University of Washington, Seattle,
Washington 98195, USA

Depurination, the spontaneous release of purine bases from DNA, occurs by the breakage of the *N*-glycosylic bond between the purine base and the deoxyribose moiety. The resulting apurinic site is relatively stable, with an *in vitro* half life of up to several days[1]. The rate constant of depurination *in vitro* has been estimated to be 1.8×10^{-9} per min (ref. 2). This corresponds to the formation of 20,000 apurinic sites per mammalian cell per day. In addition, much of the damage to DNA by chemical carcinogens occurs through covalent binding to the N-3 and N-7 positions on the purine bases[3,4], enhancing the rate of depurination of that base by three to four orders of magnitude[5–7]. The fact that cells have multiple pathways for removing apurinic sites[8,9] and even a mechanism for reinserting purines[10] suggests that the presence of apurinic sites may be highly detrimental. Apurinic sites in DNA have been postulated as potential mutagenic lesions[11], although there has been little experimental support for this concept. We have previously demonstrated that depurination of synthetic polynucleotide templates leads to increased mis-incorporation of non-complementary nucleotides[12,13]. We now present evidence that a DNA polymerase can copy past apurinic sites on natural, biologically active DNA and further, that this leads to increased mutagenesis.

Table 1 *E. coli* DNA polymerase I copies past apurinic sites on ΦX174 DNA

Fragment strand	Neutral sucrose sedimentation coefficient	Alkaline sucrose sedimentation coefficient
	Non-depurinated Z-7 fragment	
Template [^{3}H]	5.7	4.4
Product [^{32}P]	5.7	4.4
	Depurinated Z-8 fragment	
Template [^{3}H]	5.3	2.8
Product [^{32}P]	5.3	4.0

^{3}H-labelled ΦX174 single-strand *am*3 DNA was brought to a *p*H of 2.75 with HCl and incubated at 55 °C for 25 min to produce 2% depurination. The non-depurinated sample was not incubated. The samples were neutralized to *p*H 7.8 with 1 M Tris-HCl (*p*H 8.1). A fivefold molar excess of *Hae*III restriction fragment Z-1 was hybridized to the ΦX174 DNA template by first heating at 90 °C for 5 min, followed by 5 min at 0 °C and finally 30 min at 65 °C. *In vitro* synthesis was performed at 37 °C for 30 min in a reaction containing 20 mM Tris-HCl (*p*H 8.0), 2 mM dithiothreitol, 100 μM dATP, dCTP, dGTP and [α-^{32}P]dTTP (800 c.p.m. pmol^{-1}), 5 mM $MgCl_2$ and *E. coli* Pol I at a 25 : 1 molar excess to template DNA molecules. The extent of synthesis, reported as a calculated average number of nucleotides added per template molecule, was 1,370 and 4,880 for the depurinated and non-depurinated samples, respectively. The DNA was precipitated with ethanol, washed five times with 70% ethanol to remove excess salt, resuspended in 10 mM Tris-HCl (*p*H 8.0) and digested with *Hae*III restriction endonuclease as described previously[15]. The resulting restriction fragments were separated by electrophoresis on polyacrylamide gels[15], electroeluted from the gels in buffer E, ethanol precipitated and resuspended in 10 mM Tris-HCl (*p*H 8.0), 1 mM EDTA. Analysis of individual fragments by neutral and alkaline sucrose gradients was performed as previously described[15], using linear 5–20% gradients. Centrifugation at 4 °C in an SW 50.1 rotor was for 16 h at 45,000 r.p.m. Parallel gradients contained Z-5, Z-7, Z-8 and Z-10 fragments as markers.

To determine whether a DNA polymerase can copy past apurinic sites on natural DNA we characterized the product of the reaction using as a template single-stranded circular ΦX174 DNA containing a known number of apurinic sites. The experiment was performed as follows. A template strand of single-stranded ^{3}H-labelled ΦX174 *am*3 DNA was depurinated by incubation at 55 °C at a *p*H of 2.75 and a single *Hae*III restriction fragment was annealed to the template to function as a fixed primer for copying with *Escherichia coli* DNA polymerase I (Pol I) and [α-^{32}P]-labelled deoxynucleoside triphosphates (dNTPs). The product was digested with restriction endonuclease *Hae*III and the resulting restriction fragments separated by gel electrophoresis and recovered by electroelution; individual restriction fragments were analysed by sedimentation on neutral and alkaline sucrose gradients. As in our previous studies with synthetic polynucleotides[13], depurination inhibits DNA synthesis. However, three findings suggest that the enzyme copied past apurinic sites in the DNA. First, with a ΦX template (5,386 nucleotides long) containing ~50 apurinic sites, the average number of nucleotides incorporated per template was 1,370, suggesting that DNA synthesis proceeded past at least 10–15 apurinic sites (Table 1 legend). (This assumes that synthesis is synchronous and depurination is random.) Second, in this experiment synthesis was initiated with *Hae*III restriction fragment Z-1. As the fragments are synthesized in the order Z-1, Z-4, Z-8 (ref. 14), recovery of label in Z-8 indicates that synthesis proceeded past Z-4 (603 nucleotides) and Z-8 (194 nucleotides), together containing an estimated eight apurinic sites.

The third and most direct evidence that DNA polymerases copy past apurinic sites is obtained from sucrose gradient analysis of individual restriction fragments (Table 1) and takes advantage of the fact that apurinic sites in DNA are alkali labile. In neutral sucrose, both the ^{3}H-labelled template strand and the ^{32}P-labelled product strand of the isolated restriction fragment

0028-0836/81/220349—03$01.00

Table 2 Depurination decreases the accuracy of DNA synthesis *in vitro*

Apurinic sites per molecule	Nucleotides added per template	Reversion frequency $(\times10^{-6})$	Error rate
	Experiment 1: Mg^{2+}, 10× [dCTP]		
0	2,470	2.10	1/92,900
1.1	2,280	13.6	1/14,300
1.4	1,710	19.5	1/10,000
3.2	1,150	28.6	1/6,820
5.8	955	44.7	1/4,360
	Experiment 2: Mn^{2+}, equal [dNTPs]		
0	979	51.6	1/3,780
1.1	924	69.3	1/2,810
1.4	887	80.3	1/2,430
3.2	708	158	1/1,230
5.8	676	264	1/739

ΦX174 *am*3 DNA was depurinated by heating at 70 °C, *p*H 5.0. Hybridization of *Hae*III restriction fragment Z-8 and synthesis with Pol I were as described in Table 1 legend. In experiment 1, the dCTP concentration was 1 mM; in experiment 2, 0.5 mM Mn^{2+} was substituted for Mg^{2+}. The transfection assay for determination of reversion frequencies and the method for calculating the error rate are as described previously[15,16]. The background reversion frequency was 2.68×10^{-6} and 1.90×10^{-6} for experiments 1 and 2, respectively, and were subtracted from the values shown in the Table. In these experiments, the reversion frequency of uncopied DNA is the same for both depurinated and non-depurinated DNA.

show coincident peaks with either the depurinated or non-depurinated samples (Table 1). The sedimentation coefficients correspond to the length of the individual fragment under study. In alkaline sucrose, the ^3H- and ^{32}P-labelled strands obtained with a non-depurinated template remain coincident but with a sedimentation coefficient corresponding to that of a single-strand fragment. With the depurinated sample, the ^{32}P-labelled product strand sediments to the position of an intact single-strand fragment. In contrast, the ^3H-labelled template strand, which should contain alkali labile apurinic sites, sediments more slowly. This change in the sedimentation coefficient from 4.0 to 2.8 predicts two to three sites of cleavage, in agreement with our estimate of two sites in the Z-8 fragment based on 2% depurination of the ΦX template.

The mutagenic potential of apurinic sites in DNA was assessed using the recently developed ΦX assay[14–16]. The accuracy of *in vitro* DNA synthesis is quantitated by measuring the reversion frequency of an amber mutation to wild type in a transfection assay using *E. coli* spheroplasts. With ΦX174 *am*3 DNA, reversion to wild type has been shown to occur by at least two different substitutions[16]. The TAG amber codon may be converted to either a TGG or a TTG codon. In both cases, the substitution occurs opposite the adenine in the template, a potential site for depurination. Table 2 shows the effects of depurination of ΦX174 *am*3 DNA on the reversion frequency.

Because Pol I is highly accurate in Mg^{2+}-activated reactions, changes in accuracy were measured in conditions of DNA synthesis using a selective 10-fold increase in the concentration of dCTP (ref. 16), so as to obtain a baseline error rate for the non-depurinated control. With progressively depurinated templates, the reversion frequency increases from 2.1×10^{-6} with no depurination to 44.7×10^{-6} with 5.8 apurinic sites per molecule (Table 2, experiment 1). This corresponds to a change in the error rate from 1/92,900 to 1/4,360. Substitution of Mn^{2+} for Mg^{2+} results in a 100-fold increase in mis-incorporation, permitting measurement of fidelity with unbiased nucleotide pools[15,16]. With Mn^{2+}, the reversion frequency increases from 51.6×10^{-6} with a non-depurinated template to 264×10^{-6} with a template containing 5.8 apurinic sites per molecule, corresponding to a change in the error rate from 1/3,780 to 1/739.

Evidence that the enhancement of mutagenesis *in vitro* by depurination is mediated by substitution at the *am*3 locus is suggested by the specificity observed when the relative concentrations of individual nucleotides are biased. As previously reported for Pol I with non-depurinated templates in Mg^{2+}, the reversion frequency is increased in proportion to the relative concentration of dCTP in the reaction[16]. The response is linear, with a 10-fold increase in dCTP concentration giving a 10-fold increase in reversion frequency. In contrast, a 10-fold increase in dATP concentration results in only a slight increase in reversion frequency, suggesting that mis-incorporation of A (in Mg^{2+}) is normally a less frequent event at the *am*3 locus. With depurinated templates, dCTP is again mutagenic, as expected for substitution at *am*3. In addition, dATP is now highly mutagenic (Table 3), suggesting that random substitution might be occurring opposite a depurinated site in the *am*3 codon.

In our depurination conditions, the only other known lesion produced capable of reverting the amber mutation would be deamination of adenine. However, depurination is more than 10^5 times more frequent than deamination of adenine[9]. More importantly, as shown in Table 3, alkali treatment of the depurinated DNA eliminates almost all the enhanced mutagenesis resulting from depurination. Hydrolysis of the apurinic sites by alkali reduces the enhanced reversion frequencies due to depurination by 94–100%, suggesting that the mutagenic effects observed are indeed due to the presence of apurinic sites.

Even though apurinic sites on DNA have been postulated as potential mutagenic lesions[11], there has been little experimental support for this concept. The predominant biological effect of depurination previously recognized was inactivation of DNA *in vivo*[17,18]. However, depurination inhibits DNA synthesis and enhances mis-incorporation when copying either synthetic polynucleotides[12,13] or natural DNA *in vitro* (this study). As there are multiple pathways for the repair of apurinic sites *in vitro*[8–10], the likelihood of mis-coding at apurinic sites might be considerably less than that resulting from copying stable small modifications of bases such as O^6-methylguanine[19]. However, larger adducts such as those resulting from aromatic polycyclic

Table 3 Alkali treatment reverses the mutagenicity of depurination

Conditions	Non-depurinated DNA: Reversion frequency $(\times10^{-6})$	Non-depurinated DNA: Error rate	Depurinated DNA: Reversion frequency $(\times10^{-6})$	Depurinated DNA: Error rate	Alkali-treated depurinated DNA: Reversion frequency $(\times10^{-6})$	Alkali-treated depurinated DNA: Error rate	% Recovery
Equal [dNTPs]	—	1/680,000*	5.25	1/37,100	0.60	1/325,000	93.7
10× [dCTP]	2.38	1/81,900	22.7	1/8,590	3.45	1/56,500	94.7
50× [dCTP]	15.9	1/12,300	53.6	1/3,640	15.8	1/12,300	100
50× [dATP]	2.53	1/77,100	58.4	1/3,340	6.05	1/32,200	93.7

Depurination of ΦX174 *am*3 DNA was as described in Table 2 legend. Alkali treatment was performed by incubation at 37 °C for 30 min in the presence of 0.2 M NaOH followed by neutralization with 1 M HCl. Hybridization of *Hae*III restriction fragment Z-5 and synthesis conditions were as described in Table 1 legend. Nucleotide pool biases are as indicated with equal dNTPs being 50 μM. The transfection assay and determination of reversion frequency and error rate are as previously described[15,16]. The background reversion frequency, which has been subtracted from all values listed in the Table, was 1.63×10^{-6}. The survival of the depurinated DNA was 11% for uncopied DNA, representing an average of 2.2 apurinic sites per molecule[22].

* Taken from ref. 14, extrapolated from dCTP pool bias experiments.

hydrocarbons block DNA replication *in vitro*[20]. Moreover, blockage of DNA replication *in vivo* has been postulated to induce the expression of SOS functions[21]. In fact, we have been able to observe an enhanced reversion frequency when uncopied depurinated DNA is transfected in spheroplasts derived from bacteria in which error-prone DNA repair is induced by UV light[22]. Furthermore, adducts formed with aromatic hydrocarbons, particularly those involving the N-7 of guanine and the N-3 of adenine enhance depurination[5-7]. For example, 20% of the binding of the activated carcinogen, benzo[*a*]pyrene diol epoxide is on the N-7 of guanine[23]. Also, the major adduct formed with aflatoxin, a known mutagen[24] and one of the most potent carcinogens, is on the N-7 of guanine[25]. Finally, Drinkwater *et al.*[26] have recently reported a strong positive qualitative correlation between the ability of a series of carcinogens to depurinate DNA and their missense mutagenic activity. Although this correlation is not quantitative, it supports the idea that at least some of the mutagenicity associated with known carcinogens may result from depurination.

Received 8 December 1980; accepted 16 March 1981.

1. Lindahl, T. & Andersson, A. *Biochemistry* **11,** 3618–3623 (1972).
2. Lindahl, T. & Nyberg, B. *Biochemistry* **11,** 3610–3618 (1972).
3. Lawley, P. D. & Brookes, P. *Biochem. J.* **89,** 138–144 (1963).
4. Singer, B., Kroger, M. & Carrano, M. *Biochemistry* **17,** 1246–1250 (1978).
5. Margison, G. P. & O'Connor, P. J. *Biochim. biophys. Acta* **331,** 349–356 (1973).
6. Kirtikar, D. M. & Goldthwait, D. A. *Proc. natn. Acad. Sci. U.S.A.* **71,** 2022–2026 (1974).
7. Strauss, B., Scudiero, D. & Henderson, E. in *Molecular Mechanisms for Repair of DNA* Pt A (eds Hanawalt, P. C. & Setlow, R. B.) 13–24 (Plenum, New York, 1975).
8. Weiss, B. *J. biol. Chem.* **251,** 1896–1901 (1976).
9. Lindahl, T. *Prog. Nucleic Acid Res. molec. Biol.* **22,** 135–192 (1979).
10. Deutsch, W. A. & Linn, S. *Proc. natn. Acad. Sci. U.S.A.* **76,** 1089–1093 (1979).
11. Bautz, E. & Freese, E. *J. gen. Microbiol.* **22,** 113–128 (1960).
12. Shearman, C. W. & Loeb, L. A. *Nature* **270,** 537–538 (1977).
13. Shearman, C. W. & Loeb, L. A. *J. molec. Biol.* **128,** 197–218 (1979).
14. Weymouth, L. A. & Loeb, L. A. *Proc. natn. Acad. Sci. U.S.A.* **75,** 1924–1928 (1978).
15. Kunkel, T. A. & Loeb, L. A. *J. biol. Chem.* **254,** 5718–5724 (1979).
16. Kunkel, T. A. & Loeb, L. A. *J. biol. Chem.* **255,** 9961–9966 (1980).
17. Lawley, P. D. & Martin, N. C. *Biochem. J.* **145,** 85–91 (1975).
18. Bingham, P. M., Baltz, R. H., Ripley, L. A. & Drake, J. W. *Proc. natn. Acad. Sci. U.S.A.* **73,** 4159–4163 (1976).
19. Abbott, P. J. & Saffhill, R. *Biochim. biophys. Acta* **562,** 51–61 (1979).
20. Moore, P. & Strauss, B. P. *Nature* **278,** 664–666 (1979).
21. Radman, M. in *Molecular and Environmental Aspects of Mutagenesis* (eds Prakash, L., Sherman, F., Miller, M. W., Lawrence, C. W. & Taber, H. W.) 128–142 (Thomas, Springfield, 1974).
22. Schaaper, R. & Loeb, L. A. *Proc. natn. Acad. Sci. U.S.A.* **78,** 1773–1777 (1981).
23. Osborne, M. R., Harvey, R. G. & Brookes, P. *Chem.-biol. Interact.* **20,** 123–130 (1978).
24. Lin, J., Miller, J. A. & Miller, E. C. *Cancer Res.* **37,** 4430–4438 (1977).
25. Stark, A. A., Essigmann, J. M., Demain, A. L., Skopek, T. R. & Wogan, G. N. *Proc. natn. Acad. Sci. U.S.A.* **76,** 1343–1347 (1979).
26. Drinkwater, N. R., Miller, E. C. & Miller, J. A. *Biochemistry* **19,** 5087–5092 (1980).

Sym plasmid of *Rhizobium trifolii* expressed in different rhizobial species and *Agrobacterium tumefaciens*

P. J. J. Hooykaas*, A. A. N. van Brussel†, H. den Dulk-Ras*, G. M. S. van Slogteren* & R. A. Schilperoort*

Departments of Biochemistry* and Botany†, University of Leiden, The Netherlands

Bacteria of the genus *Rhizobium* fix nitrogen in symbiotic association with legumes such as clover, pea and bean. If the efficiency of this plant–bacterium relationship is to be improved or the host range of the rhizobia extended, the genes involved in the symbiosis must first be identified and localized. Accumulating genetic and biochemical data support the proposal that genes on a large plasmid[1] are involved in root nodulation, in analogy with the Ti genes of *Agrobacterium tumefaciens*, which is closely related to the fast-growing species of *Rhizobium*[2-4] (for reviews, see refs 5, 6). Although most of the data concern *Rhizobium leguminosarum*[7-11], some deal with *Rhizobium trifolii*[12-14], which associates with clover. We report here that one particular plasmid in *R. trifolii*—which we call Sym(biosis) plasmid—not only determines host specificity, but also controls several, if not most, other steps leading to root nodule formation and nitrogen fixation on clover roots.

To provide a selectable marker, we tagged the *R. trifolii* plasmid with transposon Tn*5*, which confers kanamycin resistance ($\mathrm{Km^r}$) on its host cell. *Escherichia coli* JB1830 (ref. 15), which carries an R::Mu::Tn*5* suicide plasmid coding for $\mathrm{Gm^rKm^rSm^rSp^r}$, was crossed with a $\mathrm{Rif^r}$ *R. trifolii* (LPR5001) and we then selected for $\mathrm{Km^r}$ *R. trifolii* derivatives. Several $\mathrm{Km^rRif^r}$ (but $\mathrm{Gm^sSm^sSp^s}$) transconjugants—which arose at a frequency of 4×10^{-7} per recipient—were used as possible donors in further crosses with a $\mathrm{Str^r}$ *R. trifolii* strain (LPR5020). Out of 10 possible $\mathrm{Km^r}$ donors, five transferred $\mathrm{Km^r}$ to LPR5020 at a frequency of 10^{-4}–10^{-6} per recipient. To verify that no spontaneously $\mathrm{Str^r}$ donors were present among the $\mathrm{Km^rStr^r}$ transconjugants, these were checked for and found to be $\mathrm{Rif^s}$.

To determine whether genes on the self-transmissible *R. trifolii* plasmid (referred to here as pRtr5a) that was marked with Tn*5* were involved in symbiotic nitrogen fixation, plasmid-cured derivatives were isolated as described previously[14]. After incubation of the *R. trifolii* strains for 2 days at 37 °C, we found among the surviving bacteria—most of the bacteria are killed by this heat treatment—several $\mathrm{Km^s}$ derivatives. These were $\mathrm{Nod^-}$, and by all other criteria were identical to the parental strains

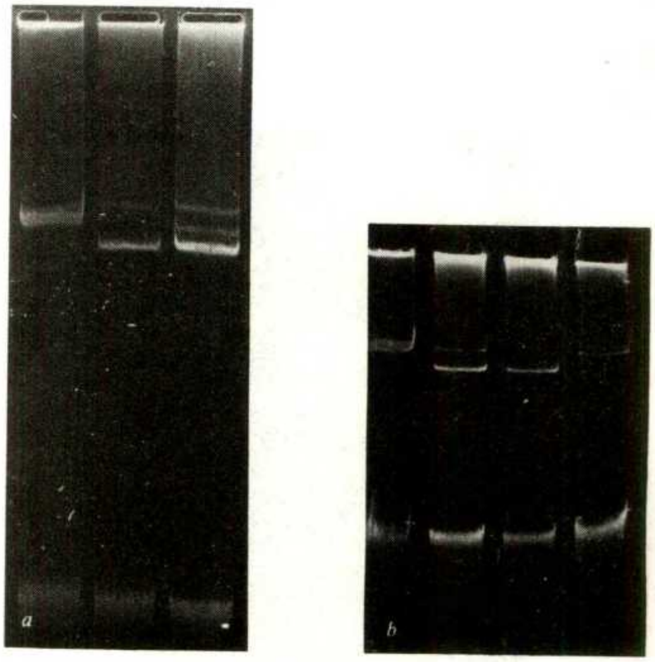


Fig. 1 Agarose gel electrophoresis of plasmid DNA isolated and run according to Casse *et al.*[22]. DNA was prepared from (from left to right): *a*: *R. trifolii* LPR5035, *R. leguminosarum* LPR1902, *R. leguminosarum* LPR1902 (pRtr5a); *b*: *R. trifolii* LPR5035, *R. meliloti* LPR2136 (pRtr5a), *R. meliloti* LPR2136, *A. tumefaciens* LBA288 (pRtr5a).

Table 1 Properties of a Sym plasmid-cured *R. trifolii* strain (LPR5039) compared with the parental strain (LPR5035)

	R. trifolii LPR5035	*R. trifolii* LPR5039
Resistance to streptomycin ($250\ \mu\mathrm{g\ ml^{-1}}$)	+	+
Growth on NB or LC medium	–	–
Sensitivity to phages LPB1, 51, 52	+	+
Uptake of Congo red from medium	–	–
Nodulation of clover seedlings*	+	–
Resistance to kanamycin ($50\ \mu\mathrm{g\ ml^{-1}}$)	+	–
Presence of pRtr5a::Tn*5*	+	–

**Trifolium pratense*.

0028–0836/81/220351—03$01.00

Table 2 Properties of *R. leguminosarum* strains before and after acquisition of the *R. trifolii* Sym plasmid

Recipient strain	Nodulation of *Trifolium pratense*		N fixation on *T. pratense*		Nodulation of *Vicia hirsuta*		N fixation on *V. hirsuta*	
	R	T	R	T	R	T	R	T
LPR1105	−	+	−	+	+	+	+	+
LPR1902	−	+	−	+	+	+	+	+
LPR1802	−	+	−	+	−	−		
LPR1120	−	+	−	+	+	+	−	

R, property of the recipient; T, property of the pRtr5a$^+$ transconjugant; LPR1105 = RCR1001 Rif^r; LPR1120 = RCR1017 Rif^r; LPR1802 = RCR1018 Rif^r; LPR1902 = RCR1039 Rif^r.

(Table 1). We showed biochemically that these strains had lost a plasmid of molecular weight 180×10^6 and that the Nod$^-$Km^s phenotype was not due to a deletion of part of the plasmid (not shown). This suggested that we had indeed marked a symbiosis-controlling (Sym) plasmid with Tn*5*. To confirm this and to exclude the possibility that secondary mutations had been introduced by the heat treatment, the plasmid was re-introduced into the Km^{s}Nod$^-$ strains by conjugation. Acquisition of pRtr5a rendered these recipients Nod$^+$Fix$^+$, showing that this self-transmissible plasmid in *R. trifolii* controls nodulation. The frequency of transfer of pRtr5a::Tn*5* to the plasmid-cured *R. trifolii* strains was as high as to the corresponding plasmid-containing strains, indicating the lack of entry exclusion functions on this plasmid. After crossing, pRtr5a$^+$ transconjugant colonies derived from the plasmid-cured *R. trifolii* strains developed more rapidly than those derived from the corresponding plasmid-containing strains. This may have been due to incompatibility between incoming and resident Sym plasmid in the latter case.

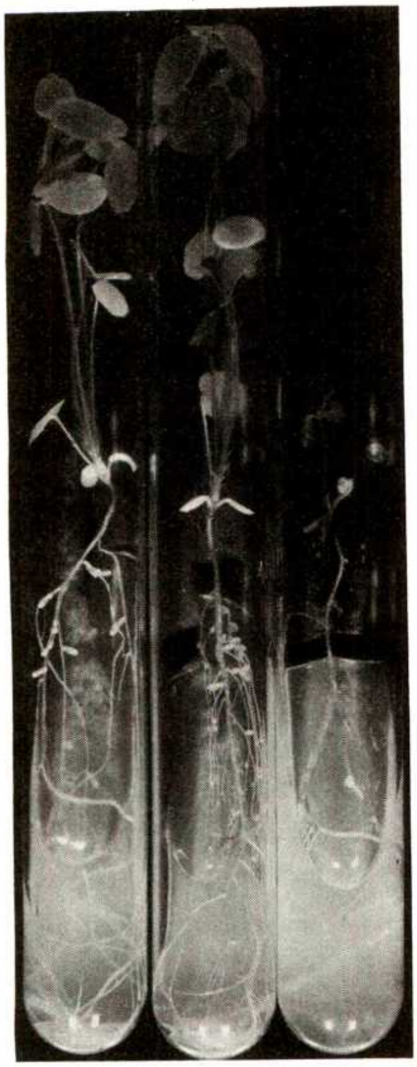


Fig. 2 Efficiency of clover plant growth in N-free medium after inoculation with: left, *R. leguminosarum* (pRtr5a::Tn*5*); centre, *R. trifolii* RCR5; right, *A. tumefaciens* (pRtr5a::Tn*5*).

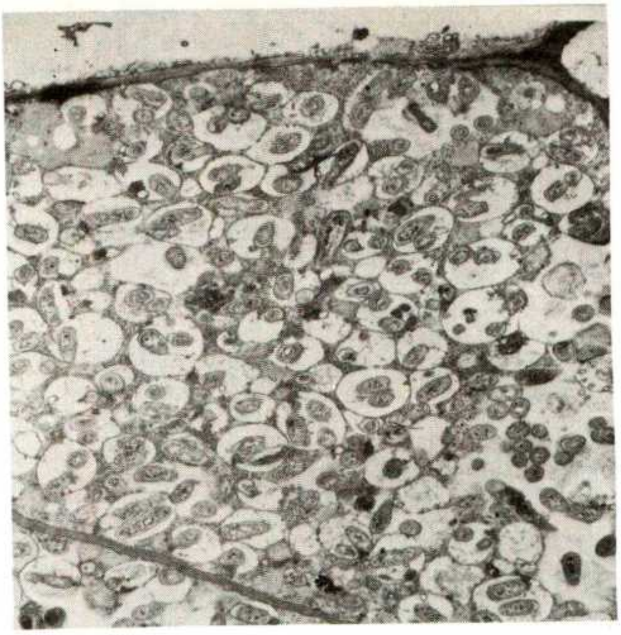


Fig. 3 Electron microscopic view inside a clover plant cell filled with agrobacteria surrounded by a membrane envelope. ×3,200.

The expression of pRtr5a was studied in several *R. leguminosarum* strains. As Table 2 shows, transfer to *R. leguminosarum*, which is closely related to *R. trifolii*[2-4] but nodulates peas and vetches, resulted in strains capable of nodulating not only their normal symbiotic partners but also clover.

Table 3 Properties of an *A. tumefaciens* transconjugant carrying pRtr5a compared with those of the parental strains

	R. trifolii‡ LPR5035	*A. tumefaciens*‡ LBA288 (pRtr5a::Tn*5*)	*A. tumefaciens*‡ LBA288
Resistance to streptomycin*	+	−	−
Resistance to rifampicin*	−	+	+
Sensitivity to phage:			
LPB 1, 51, 52	+	−	−
LPB 84	−	+	+
Growth on NB or LC medium	−	+	+
Growth on erythritol medium	+	−	−
Uptake of Congo red from medium	−	+	+
Nodulation of clover seedlings†	+	+	−
N fixation in root nodules	+	−	−
Tumour induction on plants	−	−	−

* Streptomycin at 250 μg ml^{-1}; rifampicin at 20 μg ml^{-1}.

† *Trifolium pratense*.

‡ LPR5035 = RCR5 Str^r (pRtr5a::Tn*5*); LBA288 = C58-C9 Rif^{r}Nal^r.

This suggests that the plasmid must control host specificity. The plasmid profile of a transconjugant is compared with that of the parental strains (Fig. 1). As far as Fig. 1 shows, no *R. leguminosarum* plasmid was lost from the recipients after the introduction of pRtr5a. Introduction of the plasmid into a Nod$^-$ *R. leguminosarum* mutant (LPR1802) rendered this strain Nod$^+$Fix$^+$ on clover, but it remained Nod$^-$ on peas and vetches. Plasmid pRtr5a apparently does not abolish the inability of LPR1802 to nodulate peas and vetches. Introduction of the plasmid into a Nod$^+$Fix$^-$ *R. leguminosarum* strain (LPR1120), which lacks at least a segment of the *nif* structural genes (R. K. Prakash, personal communication), rendered this strain Nod$^+$ and Fix$^+$ on clovers, showing that the Sym plasmid carries at least some of the *nif* structural genes. The latter finding was confirmed biochemically[16]. These results support the proposal that bacteria of the species *R. trifolii* and *R. leguminosarum* are identical except for their plasmid content[7].

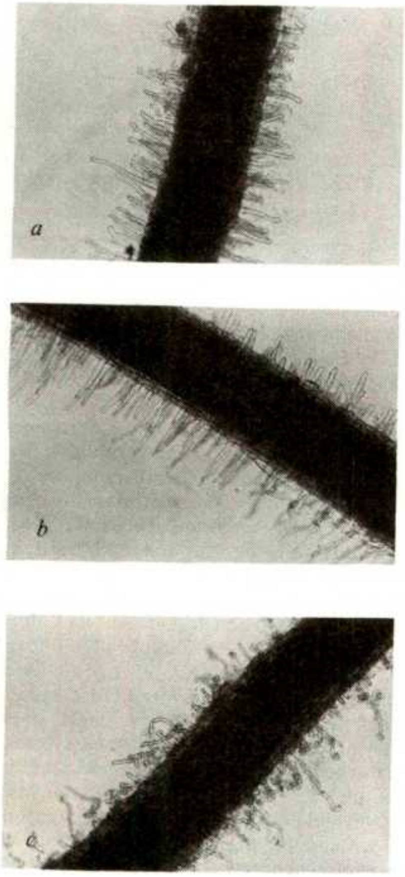


Fig. 4 Induction of root hair curling on clover roots. Clover seedlings on Jensen agar medium[23] were inoculated with bacteria, and after 5 days examined with the light microscope. *a*, No bacterial inoculation; *b*, inoculation with *A. tumefaciens* LBA288; *c*, inoculation with *A. tumefaciens* LBA288 (pRtr5a). ×180.

As we have found that the *A. tumefaciens* Ti plasmid can be transferred to *R. trifolii*, where it is stably maintained and expressed[5,17,18], we wished to see whether the *R. trifolii* Sym plasmid could be transferred to *A. tumefaciens* and whether its genes would be expressed there. To avoid interference with Ti plasmid-encoded functions, the Sym plasmid was introduced into a Ti plasmid-cured derivative of *A. tumefaciens* strain C58 (LBA288). Purified transconjugants seemed to be almost identical to the recipient LBA288 (Table 3), except that they could weakly induce white non-nitrogen-fixing root nodules on clover roots. The nodules appeared after a longer period of incubation (4–5 weeks) than did the nodules induced by *R. trifolii* (1 week). The bacteria recovered from inside the root nodules were identical to those used for inoculation, excluding the possibility that nodules had been due to contaminating rhizobia. Because of a lack of nitrogen, the plants infected with *A. tumefaciens* transconjugants did deteriorate (Fig. 2). Squashes of the *A. tumefaciens*-induced nodules examined under the light microscope revealed no typical bacteroids, such as the γ-shaped structures that were numerous in squashes of nodules induced by *R. trifolii*. Electron microscopy, however, showed that the nodules induced by *A. tumefaciens* transconjugants were true root nodules. Infection threads filled with bacteria were clearly visible. Furthermore, bacteria surrounded by a membrane were prominent inside clover plant cells, showing that the agrobacteria had been released from the infection thread into the plant cell (Fig. 3).

An essential step early in root nodule formation is the induction of a marked curling of the root hairs by the rhizobia[19]. Certain Nod$^-$ mutants cannot induce this root hair curling[20]. As our *A. tumefaciens* transconjugants with the *R. trifolii* Sym plasmid seemed to induce only sparse root nodules, we tested whether they could induce such curling. As Fig. 4 shows, there was a marked curling of clover root hairs after inoculation with the Sym plasmid harbouring *A. tumefaciens* transconjugants, but not after inoculation with *A. tumefaciens* parental strain LBA288 (Fig. 4), thus indicating that induction of root hair curling is an event controlled by the *R. trifolii* Sym plasmid. In agreement with this conclusion, *R. trifolii* strains lost the ability to induce curling of clover root hairs when cured of their Sym plasmid.

Our observation that introduction of the *R. trifolii* Sym plasmid into *A. tumefaciens* results in bacteria that, in principle, can bring about root nodulation (root hair curling, host recognition, infection thread initiation, release of membrane enclosed bacteria in the plant cell) shows that *R. trifolii* Sym plasmid can be expressed in this more distant bacterium. It is unclear why only sparse nodules, which do not fix nitrogen, appear late after infection, but the absence of typical γ-shaped bacteroids in nodules induced by *A. tumefaciens* might explain why no nitrogen is fixed in the nodules. Similar results were obtained by E. Pees and C. Wijffelman (personal communication) with root nodules induced on vetches by *A. tumefaciens* harbouring an *R. leguminosarum* Sym plasmid. These Sym plasmids are clearly different from a very large *R. trifolii* plasmid that was recently introduced into *A. tumefaciens* by mobilization with RP4 and which conferred the ability to fix nitrogen, but not the ability to induce root nodules on its new host (ref. 13 and J. Stanley, personal communication).

We thank Dr C. Wijffelman and M. P. Nuti for helpful discussions and Cebeco-Handelsraad for providing clover seeds. During preparation of this manuscript it was shown that root attachment is controlled by an *R. trifolii* plasmid[21].

Received 28 November 1980; accepted 16 March 1981.

1. Nuti, M. P., Ledeboer, A. M., Lepidi, A. A. & Schilperoort, R. A. *J. gen. Microbiol.* **100,** 241–248 (1977).
2. Heberlein, G. T., Deley, J. & Tijtgat, R. *J. Bact.* **94,** 116–124 (1967).
3. Moffet, M. L. & Colwell, R. R. *J. gen. Microbiol.* **51,** 245–266 (1968).
4. White, L. O. *J. gen. Microbiol.* **72,** 565–574 (1972).
5. Schilperoort, R. A. *et al.* in *Plasmids of Medical, Environmental and Commercial Importance* (eds Timmis, K. N. & Pühler, A.) 339–352 (Elsevier, Amsterdam, 1979).
6. Hooykaas, P. J. J., Schilperoort, R. A. & Rörsch, A. in *Genetic Engineering* Vol. 1 (eds Setlow, J. K. & Hollaender, A.) 151–179 (Plenum, New York, 1979).
7. Johnston, A. W. B. *et al. Nature* **276,** 634–636 (1978).
8. Buchanon-Wollaston, A. V. *et al. Molec. gen. Genet.* **178,** 185–190 (1980).
9. Brewin, N. J. *et al. J. gen. Microbiol.* **116,** 261–270 (1980).
10. Prakash, R. K. *et al.* in *Nitrogen Fixation* Vol. 2 (eds Newton, W. E. & Orme-Johnston, W. H.) 139–163 (University Park Press, Baltimore, 1980).
11. Nuti, M. P. *et al. Nature* **282,** 533–535 (1979).
12. Higashi, S. *J. gen. appl. Microbiol.* **13,** 391–403 (1967).
13. Stanley, J. & Dunican, L. K. *Molec. gen. Genet.* **174,** 211–220 (1979).
14. Zurkowski, W. & Lorkiewicz, Z. *Arch. Microbiol.* **123**, 195–201 (1979).
15. Beringer, J. E. *et al. Nature* **276,** 633–634 (1978).
16. Prakash, R. K., Schilperoort, R. A. & Nuti, M. P. *J. Bact.* **145,** (1981).
17. Hooykaas, P. J. J. thesis, Univ. Leiden (1979).
18. Hooykaas, P. J. J. *et al. J. gen. Microbiol.* **98,** 477–484 (1977).
19. Haack, A. *Zentbl. Bakt. Parasitkde* **117,** 343–366 (1974).
20. Dazzo, F. B. & Hubbell, D. H. *Appl. Microbiol.* **30,** 1017–1033 (1975).
21. Zurkowski, W. *Microbios* **27,** 27–32 (1980).
22. Casse, F. *et al. J. gen. Microbiol.* **113,** 229–242 (1979).
23. Jensen, H. L. *Proc. Linn. Soc. N.S.W.* **66,** 98 (1942).

Recombinant plasmids carrying nitrogen fixation genes from *Rhizobium japonicum*

Hauke Hennecke

Lehrstuhl für Mikrobiologie der Universität München,
Maria-Ward Strasse 1a, D-8000 München 19, FRG

The bacterium *Rhizobium japonicum*, which fixes atmospheric nitrogen in symbiosis with soybean roots is a member of the so-called 'slow-growing' group of *Rhizobium*, which implies that its *nif* genes are depressed in free-living, microaerobic growth conditions. This makes it a potentially useful organism for studies of the molecular biology of the regulation of symbiotic *nif* genes[1,2]. Although classical genetic experiments with this species are not well advanced, we decided to use recombinant DNA techniques to isolate and characterize its *nif* genes. Comparative studies on the biochemistry[3], immunology[2,4] and genetics[5] of nitrogenase from different N_2-fixing microorganisms suggested a fairly strong DNA sequence homology for the nitrogenase structural genes of the species investigated. We have therefore used the cloned nitrogenase structural genes from a free-living, N_2-fixing bacterium, *Klebsiella pneumoniae*[6], as hybridization probes[5] to find homologous DNA regions in a clone bank of DNA fragments of *R. japonicum*. We report here that the *nif* homology region includes two of the nitrogenase genes, *nifH* and *nifD*.

Using the restriction enzyme *Hin*dIII we have constructed a bank of 3,325 *Escherichia coli* colonies containing cloned fragments of the total genome of *R. japonicum* strain 110 (ref. 7). The cloning vector was pBR322 (ref. 8), which has a single *Hin*dIII site within the genes coding for tetracycline resistance. The *K. pneumoniae* DNA probe used for hybridization came from plasmid pSA30 (ref. 6), and represents an *Eco*RI fragment containing almost exclusively the nitrogenase structural genes, *nifK*, *nifD* and *nifH* (Fig. 1). This *Klebsiella* DNA fragment was purified by preparative agarose gel electrophoresis and nick-translated[9] in the presence of ^{32}P-thymidine-triphosphate to high specific radioactivity (10^8 c.p.m. per μg DNA). Colonies of the colony bank were grown on nitrocellulose filters to a size of ~2–3 mm. The filters were then layered on to LB agar[16] containing 100 μg ml^{-1} chloramphenicol (for 24 h) to facilitate amplification of the recombinant plasmids within the cells.

The ^{32}P-labelled *Klebsiella nif* DNA was used to detect *R. japonicum nif* genes by colony hybridization[10]; of the 3,325 colonies screened, 21 contained DNA that clearly hybridized. The recombinant plasmids were isolated from all 21 strains. Each plasmid had a *Hin*dIII DNA insert of molecular weight (MW) 4.85×10^6; five of these had a second insert and one had two further inserts (not shown). Southern blotting experiments[11] showed that in all 21 cases it was the 4.85×10^6 MW fragment that hybridized to *Klebsiella nifKDH* DNA. Plasmid pRJ676 from colony 676, a representative of the first group of recombinant plasmids, was studied in more detail.

First, a detailed restriction map of pRJ676 was established, using the restriction enzymes *Hin*dIII, *Eco*RI, *Pst*I, *Bam*HI and *Sal*I (Fig. 1). Single digests as well as double digests in every possible combination were done. The molecular weights of the fragments were determined using molecular weight markers (digested phage λ DNA), and we used the known restriction sites on pBR322 (ref. 12) to determine the molecular weights of those fragments adjacent to the vector. Precise location of the restriction sites was made possible by isolation and analysis of nine individual subclones from the various fragments of the 4.85×10^6 MW pRJ676 insert (Fig. 1). The vectors used for recloning the smaller DNA pieces are described in Fig. 1 legend. From this restriction map it was possible to determine the approximate location of the DNA region within pRJ676 that exhibits homology with *K. pneumoniae nifKDH* DNA. These experiments were done by transfering restriction patterns of pRJ676 from the agarose gel on to nitrocellulose filters (Southern blotting[11]), and hybridizing them with the nick-translated *Eco*RI insert of pSA30. A few examples are shown in Fig. 2*A*, for which we used the *Hin*dIII, *Bam*HI, *Hin*dIII/*Bam*HI and *Eco*RI restriction patterns. The hybridizing band could be clearly identified by direct comparison with the stained bands on the gels.

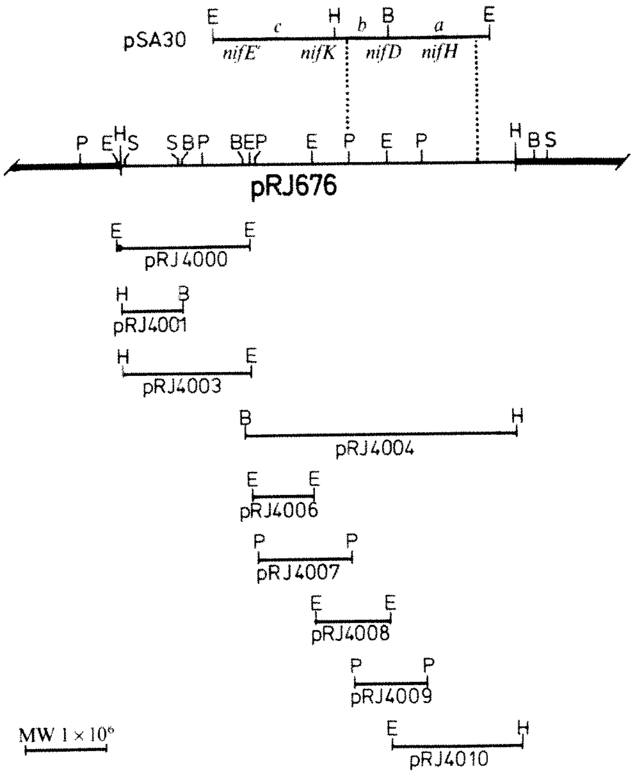


Fig. 1 Recombinant plasmids used in this study. For all plasmids, except pRJ676, only the cloned DNA inserts are shown. Restriction sites are: E, *Eco*RI; B, *Bam*HI; H, *Hin*dIII; P, *Pst*I; S, *Sal*I. The scale bar represents a length of DNA of MW 1×10^6. The *Eco*RI fragment of pSA30 (ref. 6) is shown together with the approximate location of the *K. pneumoniae* nitrogenase structural genes, *nifK D* and *H*, and with the subfragments *a*, *b* and *c*, which were used for the hybridization experiment in Fig. 2*B*. pRJ676 is one representative of the recombinant plasmids carrying the *nif* homologous region of *R. japonicum*. The cloning vector pBR322 (ref. 8) is indicated by a heavy line, and is arbitrarily split into two halves. The dotted lines enclose the approximate region of DNA homology between the *K. pneumoniae* and *R. japonicum nif* DNAs. A series of restriction fragments were cloned into the following cloning vectors (in parentheses): pRJ4006 (pACYC184, ref. 14); pRJ4000, pRJ4008 (pUR2, ref. 15); all others (pBR322, ref. 8). In these subcloning experiments pRJ676 and the respective cloning vectors were digested with one single or two different restriction enzymes, and the resulting fragments ligated with DNA ligase. The rejoined DNA was transformed into *E. coli* strain HB101 (when pBR322 or pACYC184 were the vectors) or into strain 79-02 (when pUR2 was the vector). The transformants were tested for the presence (or absence, due to insertion) of the vector's antibiotic resistance phenotype(s). In the case of vector pUR2, clones with inserted foreign DNA could be visualized directly as white colonies on Xgal plates[15]. Plasmid DNA of presumptive subclones was then isolated and subjected to DNA restriction enzyme analysis.

Next we asked which of the structural genes of the *K. pneumoniae* nitrogenase complex—*nifH*, *nifD* or *nifK*— are homologous to *R. japonicum nif* DNA. To find the answer we purified DNA subfragments of the *Klebsiella nifKDH* region (designated fragments *a*, *b* and *c* in Fig. 1) as molecular hybridization probes for the *Eco*RI pattern of pRJ676. Figure 2*B*

0028-0836/81/220354—02$01.00

shows that subfragment *a* (containing *nifH* and part of *nifD*) hybridizes to the large *Eco*RI fragment of pRJ676, and subfragment *b* (containing part of *nifD* and part of *nifK*) hybridizes to one of the two smaller *Eco*RI fragments of pRJ676. Subfragment *c* (containing the larger part of *nifK* and part of the genes downstream of the *nif* cluster) seems to have no DNA homology with any region of pRJ676. Thus, interspecies hybridization seems to be based mainly on *nifH* and *nifD* homologies. However, these experiments do not completely exclude the possibility that at least part of *nifK* DNA also hybridized to *R. japonicum* DNA.

Several control experiments confirmed these results: (1) the cloned *R. japonicum nif* fragment of MW 4.85×10^6 was itself purified and used as a hybridization probe on the *Eco*RI/*Bam*HI/*Hin*dIII pattern of pSA30—only the *Klebsiella* subfragments *a* and *b* hybridized. (2) Vector pBR322 was used as a hybridization probe for restriction fragments of pRJ676, and only DNA fragments containing parts of pBR322 hybridized. (3) The *Eco*RI fragment of pSA30 was used to probe some of the individual subclones of pRJ676, verifying the approximate region of homology described here. (4) The 4.85×10^6 MW *R. japonicum* DNA was found to hybridize to plasmid pGR151, which contains a cloned *Klebsiella* DNA fragment coding for the amino-terminal part of the *nifH* gene product (G. E. Riedel, unpublished results). This supports the finding that pRJ676 is homologous to *nifH* of *K. pneumoniae*.

Thus, Fig. 1 shows the approximate region of homology between the *K. pneumoniae* and *R. japonicum nif* DNAs. Ruvkun and Ausubel[5] reported the sucessful cloning of part of the *nif* genes from *R. meliloti*. Their results agree well with ours, and it is intriguing that only two genes out of the 17 *Klebsiella nif* genes share apparent DNA sequence homology. It will be interesting to determine whether the cloned DNA regions of pRJ676 adjacent to the *nif*-homology region code for proteins of the symbiotic nitrogen-fixing apparatus. The availability of cloned *R. japonicum nif* DNA in *E. coli* makes it possible to alter this region genetically, and preliminary experiments have shown that transposable DNA elements[13] can be introduced easily into the cloned *R. japonicum nif* DNA (our unpublished results).

This work was sponsored by a grant from the Deutsche Forschungsgemeinschaft. I thank Dr A. Böck for support, Drs F. Ausubel, F. Cannon, G. Riedel, G. Ruvkun and U. Rüther for bacterial strains and helpful suggestions, and P. Frischeisen, M. Fuhrmann and K. Kaluza for technical assistance.

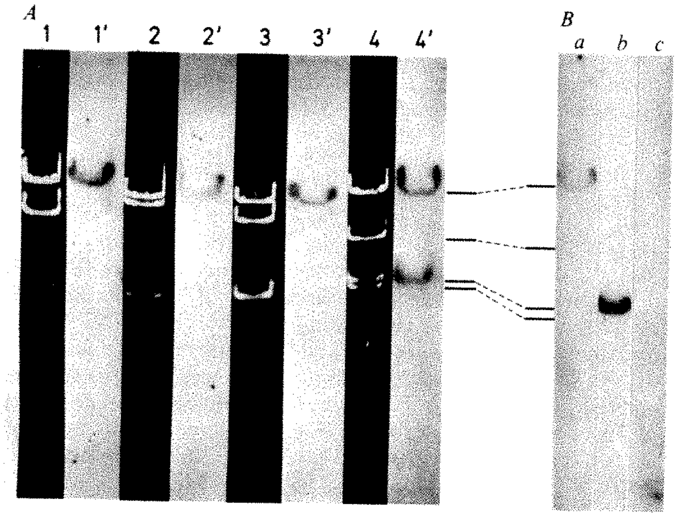


Fig. 2 Hybridization of DNA fragments of pRJ676 with ^{32}P-labelled *K. pneumoniae nif* DNA. *A*, Restriction fragments were generated by digestion with the enzymes *Hin*dIII (1), *Bam*HI (2), *Hin*dIII/*Bam*HI (3) and *Eco*RI (4). They were then separated on 0.8% agarose gels and stained with ethidium bromide. Duplicates were transferred on to nitrocellulose filters and hybridized with the nick-translated *Eco*RI fragment of pSA30. The white strips (1′, 2′, 3′ and 4′) show the respective autoradiograms. *B*, *Eco*RI fragments of pRJ676 were hybridized in three separate experiments with nick-translated subfragments *a*, *b* and *c* of pSA30 (compare with Fig. 1). Only the respective autoradiogram strips are shown.

Received 21 November 1980; accepted 6 March 1981.

1. Lim, S. T., Hennecke, H. & Scott, D. B. *J. Bact.* **139,** 256–263 (1979).
2. Scott, D. B., Hennecke, H. & Lim, S. T. *Biochim. biophys. Acta* **565,** 365–378 (1979).
3. Emerich, D. W. & Burris, R. H. *J. Bact.* **134,** 936–943 (1978).
4. Maier, R. J. & Brill, W. J. *J. Bact.* **127,** 763–769 (1976).
5. Ruvkun, G. B. & Ausubel, F. M. *Proc. natn. Acad. Sci. U.S.A.* **77,** 191–194 (1980).
6. Cannon, F. C., Riedel, G. E. & Ausubel, F. M. *Molec. gen. Genet* **174,** 59–66 (1979).
7. Hennecke, H. & Mielenz, J. R. in *Genetic Engineering for Symbiotic Nitrogen Fixation and Conservation of Nitrogen* (eds Rains D. W. *et al.*) (Plenum, New York, in the press).
8. Bolivar, F. *et al. Gene* **2,** 95–113 (1977).
9. Maniatis, T., Jeffrey, A. & Kleid, D. G. *Proc. natn. Acad. Sci. U.S.A.* **72,** 1180–1188 (1975).
10. Grunstein, M. & Hogness, D. S. *Proc. natn. Acad. Sci. U.S.A.* **72,** 3961–3965 (1975).
11. Southern, E. M. *J. molec. Biol.* **98,** 503–517 (1975).
12. Sutcliffe, J. G. *Cold Spring Harb. Symp. quant. Biol.* **43,** 77–90 (1978).
13. Kleckner, N., Roth, J. & Botstein, D. *J. molec. Biol.* **116,** 125–159 (1977).
14. Chang, A. C. Y. & Cohen, S. N. *J. Bact.* **134,** 1141–1156 (1978).
15. Rüther, U. *Molec. gen. Genet.* **178,** 475–477 (1980).
16. Miller, J. H. *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory, New York, 1972).

Synthesis and characterization of (2′-5′)ppp3′dA(p3′dA)$_n$, an analogue of (2′-5′)pppA(pA)$_n$

Paul Doetsch, Joseph M. Wu*, Yosuke Sawada & Robert J. Suhadolnik†

Department of Biochemistry, Temple University School of Medicine, Philadelphia, Pennsylvania 19140, USA

The secretion of interferon after viral infection enhances antiviral resistance in other cells characterized by increased activity of certain enzymes, a protein kinase, an endoribonuclease and a (2′-5′) (A)$_n$ synthetase[1–6]. The biochemical function of the synthetase is to convert ATP into oligonucleotides called generically (2′-5′)pppA(pA)$_n$, which are characterized by the 2′,5′-phosphodiester bond (see Table 1 legend for abbreviations used). The function of these (2′-5′) oligonucleotides seems to be the activation of the latent endoribonuclease whose consequence is to inhibit protein synthesis. As part of our study of the development of the antiviral/antitumour state, we have set out to synthesize analogues of (2′-5′)pppA(pA)$_n$ in the expectation that these analogues would also activate a latent endoribonuclease, yet be metabolically more stable. We now describe the enzymatic conversion of 3′dATP (cordycepin 5′-triphosphate) to the 5′-triphosphate trimer, (2′-5′)ppp3′dA(p3′dA)$_2$, the effectiveness of the latter in the inhibition of translation and its resistance to hydrolysis by 2′,5′-phosphodiesterase[7]. The enzymatic synthesis of the 2′,5′-phosphodiester bond with 3′dATP by (2′-5′) (A)$_n$ synthetase is an extension of our earlier report on the formation of a 2′,5′-phosphodiester bond following the incorporation of 3′dAMP from [G-^{3}H]cordycepin into the internucleotide linkage of RNA of H.Ep. 1 cells[8].

The per cent conversions of ATP and 3′dATP to (2′-5′) oligonucleotides following 5- and 17-h incubations are shown in Table 1. The radioactive oligonucleotides synthesized following incubations with either ATP or 3′dATP eluted from the poly(rI)·poly(rC)–agarose columns were diluted to 50 mM KCl and analysed on DEAE-cellulose in a 7 M urea–NaCl gradient with authentic (2′-5′)pppA(pA)$_2$ (P-L Biochemicals). Six radioactive compounds were detected following incubations with [G-^{3}H]3′dATP (Fig. 1*a*). One, with a charge of −6, was displaced together with authentic marker (2′-5′)pppA(pA)$_2$ (peak tube at 36 ml, Fig. 1*a*). The ^{3}H-nucleotide, with a charge

---

* Present address: Department of Biochemistry, New York Medical College, Valhalla, New York 10595, USA

† To whom correspondence should be addressed.

0028–0836/81/220355—04$01.00

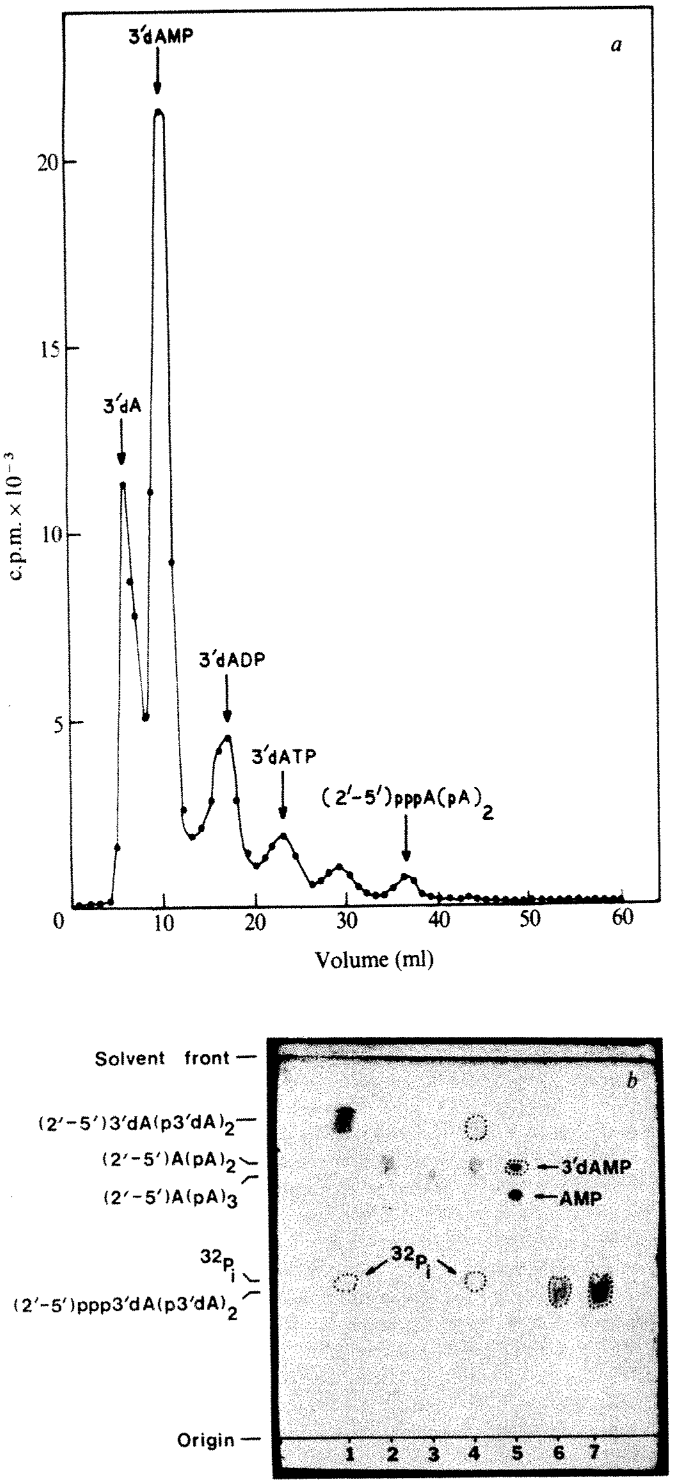


Fig. 1 *a*, Chromatography of the analogue, ^{3}H-(2'-5')ppp3'dA(p3'dA)$_n$, from incubations with [G-^{3}H]3'dATP on DEAE-cellulose. Peak fractions of material washed from poly(rI)·poly(rC)-agarose columns from a 17-h incubation with [G-^{3}H]3'dATP were applied to a DEAE-cellulose column (Whatman DE52, 0.5 × 17 cm). The nucleotides were displaced with a 50–150 mM linear gradient of NaCl (40 ml/40 ml) and 50 mM Tris-HCl, *p*H 8.0, in 7 M urea; 1-ml fractions were collected; flow rate 4 ml h^{-1}. Cordycepin (3'dA), 3'dAMP, 3'dADP, 3'dATP and authentic (2'-5')pppA(pA)$_2$ were included as markers as indicated by arrows and were monitored on a recording spectrophotometer (LKB 2138 Uvicord S). The elution profile represents a 0.10-ml aliquot of each 1-ml fraction assayed for radioactivity. A similar elution profile was obtained from a 17-h incubation with [α-^{32}P]3'dATP (data not shown). In the other incubations with [α-^{32}P]-3'dATP and [8-^{3}H]ATP, the radioactive material eluting from the poly(rI)·poly(rC)–agarose columns was chromatographed on DEAE-cellulose columns (0.6 × 2.0 cm) and washed with 42 ml of buffer B: 0.09 M KCl, 20 mM HEPES, *p*H 7.5. The oligonucleotides were displaced with 5 ml of buffer B supplemented with 0.35 M KCl. The radioactive nucleotides in the 0.35 M KCl were further fractionated by the 7 M urea–NaCl gradient–DEAE-cellulose chromatography method described above. *b*, Photograph taken under UV light of cellulose TLC (Eastman no. 13254) developed in solvent A. Dark areas indicate authentic (chemically synthesized) nucleotides detected by UV light; dotted circles indicate areas of ^{32}P. Following photography, lanes 1–7 from the origin to the solvent front were cut into 1-cm^2 pieces, eluted with water and assayed for radioactivity in 5 ml aqueous counting scintillant (Amersham). Lane 1: authentic unlabelled core (2'-5')-3'dA(p3'dA)$_2$ (R_F 0.81) detected by UV light and ^{32}P$_i$ (standard, R_F 0.41) indicated by dotted circle. Lanes 2, 3: authentic unlabelled core (2'-5')-A(pA)$_2$ (R_F 0.71) and authentic unlabelled core (2'-5')A(pA)$_3$ (R_F 0.68), respectively, detected by UV light. Lane 4: bacterial alkaline phosphatase digest of ^{32}P-(2'-5')ppp3'dA(p3'dA)$_2$; core ^{32}P-(2'-5')3'dA(p3'dA)$_2$ (R_F 0.81, 2,397 d.p.m.) and ^{32}P$_i$ (R_F 0.40, 2,829 d.p.m.) indicated by dotted circles (digestion did not go to completion; when digestion of similar preparations did go to completion, 34% of the ^{32}P was isolated as ^{32}P$_i$ by XAD-4 chromatography, see above); authentic unlabelled core (2'-5')-3'dA(p3'dA)$_2$ (R_F 0.81), authentic unlabelled core (2'-5')A(pA)$_2$ (R_F 0.71) and authentic unlabelled core (2'-5')A(pA)$_3$ (R_F 0.68) detected by UV light (no radioactivity was detected in the core (2'-5')- A(pA)$_2$ or core (2'-5')-A(pA)$_3$ regions). Lane 5: snake venom phosphodiesterase digest of core ^{32}P-(2'-5')3'dA(p3'dA)$_2$; ^{32}P-3'dAMP (R_F 0.69, 1,228 d.p.m.) indicated by dotted circles; authentic 3'dAMP (R_F 0.69) and AMP (R_F 0.62) detected by UV light (no radioactivity was detected in the AMP region). Lane 6: ^{32}P-(2'-5')ppp3'dA(p3'dA)$_2$ (R_F 0.36, 1,729 d.p.m.) indicated by dotted circle and authentic (2'-5')pppA(pA)$_2$ (R_F 0.36) detected by UV light. Lane 7: T2 RNase digest of ^{32}P-(2'-5')ppp3'dA(p3'dA)$_2$; the ^{32}P-(2'-5')-ppp3'dA(p3'dA)$_2$ (R_F 0.36, 1,722 d.p.m.) indicated by dotted circle co-chromatographed with authentic marker (2'-5')pppA(pA)$_2$ (R_F 0.36 detected by UV light). In control experiments, 0.5% core ^{32}P-(2'-5')A-(pA)$_3$ could be detected when mixed with 99.5% core ^{32}P-(2'-5')3'dA-(p3'dA)$_2$ and chromatographed (Whatman 3MM, solvent A).

of −6, has been identified as the triphosphate trimer analogue of 3'dATP, that is, (2'-5')ppp3'dA(p3'dA)$_2$.

Several experiments were done to characterize the structure of the putative (2'-5')ppp3'dA(p3'dA)$_2$ and to establish its biological properties. Treatment of the putative ^{32}P-(2'-5')ppp3'dA(p3'dA)$_2$ with bacterial alkaline phosphatase (BAP) followed by XAD-4 chromatography[9] resulted in the isolation of 34% of the ^{32}P as ^{32}P$_i$ (theoretical, 33.3%). Additional proof for the structure of (2'-5')ppp3'dA(p3'dA)$_2$ was obtained by treatment of the adenylate analogue with BAP, snake venom phosphodiesterase I (SVPD I) and T2 RNase followed by TLC and measurement of radioactive compounds (Fig. 1*b*). The nucleotide synthesized from [α-^{32}P]3'dATP and purified by DEAE-cellulose chromatography (Fig. 1*a*, charge −6) showed only one radioactive spot (Fig. 1*b*, lane 6). BAP hydrolysis[10] of the putative [α-^{32}P](2'-5')ppp3'dA(p3'dA)$_2$ produced ^{32}P$_i$ and core ^{32}P-(2'-5')3'dA(p3'dA)$_2$ which had the same R_F as authentic, chemically pure core 3'dA(p3'dA)$_2$ (ref. 11) (Fig. 1*b*, lane 4). There was no ^{32}P in the region of (2'-5')A(pA)$_3$ (lane 4), indicating that putative ^{32}P-(2'-5')ppp3'dA(p3'dA)$_2$ was not contaminated with (2'-5')ppA(pA)$_3$ (charge −6). Hydrolysis of core ^{32}P-3'dA(p3'dA)$_2$ with SVPD I[10] formed ^{32}P-3'dAMP (Fig. 1*b*, lane 5); there was no ^{32}P in the AMP region. Therefore, the [α-^{32}P]3'dATP could not be contaminated with [α-^{32}P]ATP. (Of eight commercial preparations of [α-^{32}P]3'dATP, only two samples were free of contamination with [α-^{32}P]ATP. The remaining six preparations showed ^{32}P in the regions equivalent to ATP or AMP by paper chromatography and TLC (solvents A and B).) Further proof of the 2', 5' linkage was obtained by the absence of hydrolysis (Fig. 1*b*, lane 7) on treatment of the putative (2'-5')ppp3'dA(p3'dA)$_2$ with T2 RNase[10]. Core ^{32}P-(2'-5')3'dA(p3'dA)$_2$, ^{32}P-3'dAMP and ^{32}P-(2'-5')ppp3'dA(p3'dA)$_2$ (Fig. 1*b*, lanes 4–6, respectively) following elution and rechromatography on cellulose TLC plates in solvent B showed only one radioactive spot each which co-chromatographed with their respective standards.

The inhibition of translation by the analogue (2'-5')ppp3'dA(p3'dA)$_2$ was compared with the inhibition by tetramer adenylate (2'-5')pppA(pA)$_3$ (Fig. 2*a*) and the trimer adenylate (2'-5')pppA(pA)$_2$ (Fig. 2*b*). The inhibition of translation by 40 nM (2'-5')pppA(pA)$_3$ was 52% and decreased to control values at 0.004 nM; that by 22 nM (2'-5')ppp3'dA(p3'dA)$_2$ was 68% and decreased to control

Table 1 Enzymatic synthesis of (2'-5')ppp3'dA(p3'dA)$_n$ and (2'-5')pppA(pA)$_n$

Substrate	Final concentration (mM)	Incubation time (h)	Conversion to (2'-5')oligo-nucleotide (%)
[8-^{3}H]ATP	2	5	21.9
[8-^{3}H]ATP	2	17	19.0
[α-^{32}P]3'dATP	2	5	0.3
[α-^{32}P]3'dATP	2	17	3.0
[G-^{3}H]3'dATP	1.3	17	3.7

Poly(rI) · poly(rC)–agarose columns (0.5 × 1.5 cm) bound with (2'-5')(A)$_n$ synthetase were prepared as described previously[19], except that reticulocyte lysate (1.2 ml) was used in place of HeLa cell extract and the column wash buffer consisted of 100 mM KCl, 2 mM Mg(OAc)$_2$, 2 mM dithiothreitol, 10% glycerol and 20 mM HEPES, *p*H 7.5 (buffer A). The final concentration of Mg(OAc)$_2$ was 8 mM in all incubations except for the [G-^{3}H]3'dATP experiment where the concentration was 2.3 mM. Columns were incubated at 30 °C for the time periods indicated with [8-^{3}H]ATP (10 μCi, 22 Ci mmol^{-1}), [α-^{32}P]3'dATP (25 μCi, 640 or 3,000 Ci mmol^{-1}, Amersham, NEN) or [G-^{3}H]3'dATP (10 μCi, 640 mCi mmol^{-1}). Ninety-five per cent of the radioactivity was eluted with 2 ml of buffer A. Cordycepin (3'-deoxyadenosine) was isolated and purified from *Cordyceps militaris* (m.p. 228–229°C, literature 231°C)[20]. Commercially available 3'dATP of the highest purity was purchased (from Sigma) or synthesized in this laboratory[21]. [G-^{3}H]3'dATP was synthesized in this laboratory, oxidized with periodate to remove contaminating ATP or other ribonucleotides and repurified by descending paper chromatography (Whatman 3MM in solvents A and B; solvent A: isobutyric acid/concentrated ammonia/water, 66:1:33, v/v/v; solvent B: 1-propanol/concentrated ammonia/water, 60:30:10, v/v/v). In solvents A and B, one UV light absorbing region or one radioactive region was detected for 3'dATP, [G-^{3}H]3'dATP or [α-^{32}P]3'dATP. The ratio of tritium in the adenine:3-deoxyribose of the [G-^{3}H]cordycepin was 9:1 (ref. 21). Product formation was determined by the amount of radioactivity displaced from the DEAE-cellulose column with 0.35 M KCl divided by the total radioactivity recovered, except for the [G-^{3}H]3'dATP experiment. Per cent conversion for the [G-^{3}H]3'dATP experiment was determined from the elution profile shown in Fig. 1*a* as the percentage of radioactivity eluting after 3'dATP because chromatography of this material on DEAE-cellulose and elution with 0.35 M KCl was not performed before the DEAE-cellulose chromatography NaCl gradient elution method shown in Fig. 1*a*. The amount of ^{32}P displaced with 0.35 M KCl when 25μCi of [α-^{32}P]3'dATP was added to DEAE-cellulose columns was 0.15%, which was subtracted from the ^{32}P in the 0.35 M KCl fraction. Abbreviations used are: A, adenosine; 3'dA, 3'-deoxyadenosine (cordycepin); pA, 5'AMP; p3'dA, 3'-deoxy-5'-AMP; 3'dATP, 3'-deoxyATP; (2'-5')pppA(pA)$_2$, (2'-5')pppA(pA)$_3$, or (2'-5')ppp3'dA-(p3'dA)$_2$, the 2',5'-linked trimers or tetramer of either AMP or 3'dAMP with a triphosphate at the 5' end; core, (2'-5')A(pA)$_n$ or (2'-5')3'dA(p3'dA)$_n$.

values at 0.04 nM (Fig. 2*a*). The trimer analogue (2'-5')ppp3'dA-(p3'dA)$_2$ is about four times more potent an inhibitor of translation than is (2'-5')pppA(pA)$_2$ (Fig. 2*b*). The inhibition of translation by (2'-5')ppp3'dA(p3'dA)$_2$ ± periodate did not change, although there was a small change in the inhibition of translation by periodate-treated (2'-5')pppA(pA)$_2$ (Fig. 2*b*). These data indicate that there is no ribosyl moiety on the 3' terminus of (2'-5')ppp3'dA(p3'dA)$_2$. In comparing the inhibition of translation by (2'-5')ppp3'dA(p3'dA)$_2$ with that of (2'-5')-pppA(pA)$_2$ (Fig. 2*b*), it must be kept in mind that Williams *et al.*[12] reported that in the lysed rabbit reticulocyte system (2'-5')pppA(pA)$_2$ is not a good inhibitor of translation. The translation experiments show inhibition of translation by 0.4 nM analogue (2'-5')ppp3'dA(p3'dA)$_2$, but no inhibition by 0.004 nM (2'-5')pppA(pA)$_3$; thus, any possible contamination of [α-^{32}P]3'dATP with [α-^{32}P]ATP must be <1%.

The stability of (2'-5')ppp3'dA(p3'dA)$_n$ in HeLa cell extracts was compared with that of (2'-5')pppA(pA)$_n$. There was no hydrolysis of the analogue (2'-5')ppp3'dA(p3'dA)$_n$ (12 μM) by HeLa cell extracts[13] following a 45-min incubation; however, the $t_{1/2}$ of (2'-5')pppA(pA)$_n$ (5 μM) is about 2 min (Fig. 3). With 500 μM (2'-5')pppA(pA)$_n$, there was a 50% hydrolysis in 45 min as determined by DEAE-cellulose chromatography; with 500 μM authentic core analogue (2'-5')3'dA(p3'dA)$_2$ there was no hydrolysis as determined by cellulose TLC (solvents A and B). Neither cordycepin nor 3'dAMP was detected. The lack of hydrolysis of the analogue may be similar to the lack of hydrolysis of (3'-5')2'-deoxyribonucleotides by the 2', 5'-phosphodiesterase[14].

Thus, the data presented show that (1) the analogue of (2'-5')pppA(pA)$_n$ can be synthesized enzymatically from 3'dATP by the (2'-5')(A)$_n$ synthetase from lysed rabbit reticulocytes, (2) the 3'-hydroxyl on the ribose moiety of ATP

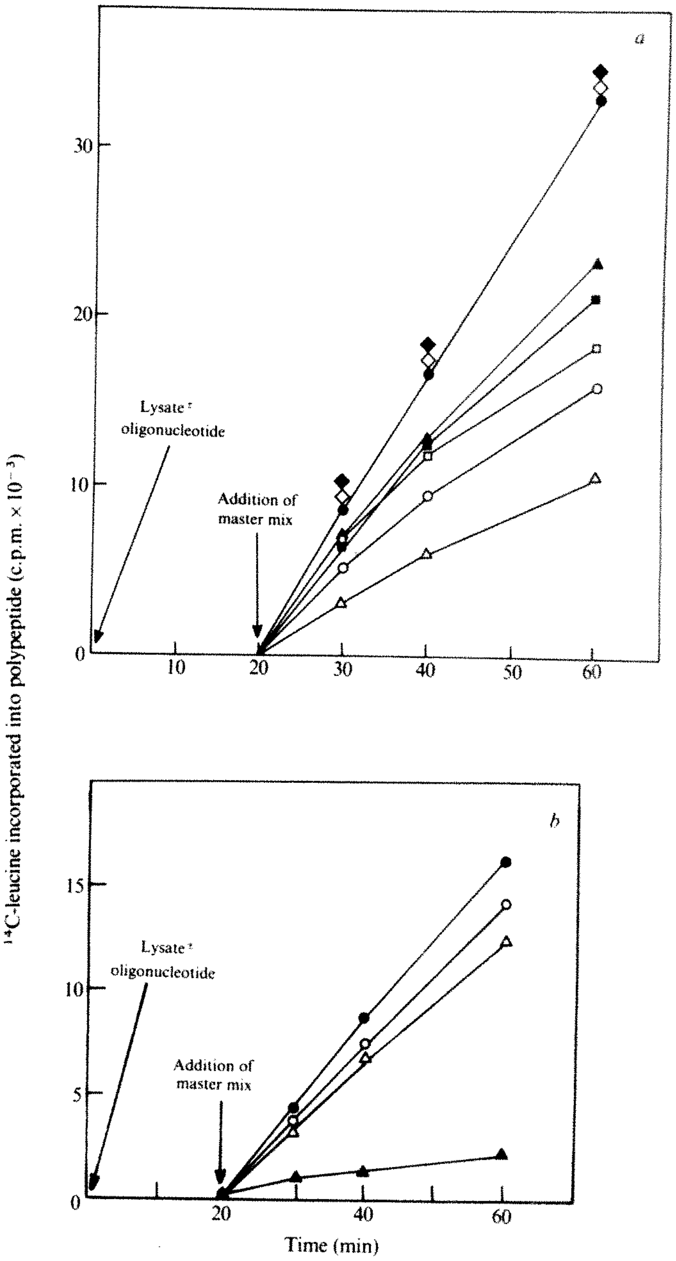


Fig. 2 *a*, Inhibition of translation by (2'-5')pppA(pA)$_3$ (tetramer adenylate) at 40 nM (○), 0.4 nM (□), 0.04 nM (■) and 0.004 nM (◆) were compared with trimer 3'-deoxyadenylate at 22 nM (△), 0.4 nM (▲), 0.004 nM (◇) and control (●). Concentrations are expressed in terms of AMP and 3'dAMP equivalents. (2'-5')pppA(pA)$_3$ was synthesized and isolated as described elsewhere[23]. (2'-5')ppp3'dA(p3'dA)$_2$ was obtained by DEAE-cellulose chromatography (Fig. 1*a*, marker (2'-5')pppA(pA)$_2$ was omitted in this purification). The fractions corresponding to putative ^{32}P-(2'-5')ppp3'dA(p3'dA)$_2$ from the NaCl gradient elution method were pooled, desalted by gel filtration (Sephadex G-10, 0.9 × 87 cm), and concentrated *in vacuo*. Translation was assayed by the incorporation of L[U-^{14}C]leucine (0.60 μCi, 335 mCi mmol^{-1}) into trichloroacetic acid-insoluble polypeptides. The data were obtained by using 5-μl aliquots of the assay volume (30 μl) processed as described previously[24]. Oligonucleotide samples in 3.0 μl were preincubated in the rabbit reticulocyte lysate (15 μl, Clinical Convenience Inc.) for 20 min before the start of translation[24]. Master mix components and their final concentrations in the lysate were as follows: ATP, 1.0 mM; GTP, 0.2 mM; phosphocreatine, 15 mM; creatine phosphokinase, 1.3 units per 30 μl; KCl, 85 mM; haemin, 25 μM. *b*, Inhibition of translation by (2'-5')pppA(pA)$_2$ (trimer adenylate) and (2'-5')-ppp3'dA(p3'dA)$_2$ (analogue trimer 3'-deoxyadenylate). Inhibition of translation with (2'-5')pppA(pA)$_2$ (100 nM) (△) and 100 nM plus periodate (○) were compared with that of (2'-5')ppp3'dA(p3'dA)$_2$ (100 nM) or 100 nM plus periodate (▲) and control (●). (2'-5')ppp3'dA(p3'dA)$_2$ was isolated as described above. Translation assays were performed as described above except that 0.25 μCi of L-[U-^{14}C]leucine (355 mCi mmol^{-1}) was used and the assay volume was 50 μl. Oligonucleotide samples (5 pmol) were treated with 1.25 nmol of sodium metaperiodate[22] in a final volume of 10 μl for 20 min at 25 °C in the dark. Periodate-treated oligonucleotides were tested in translation assays as described above. Periodate in samples containing no oligonucleotide did not inhibit translation.

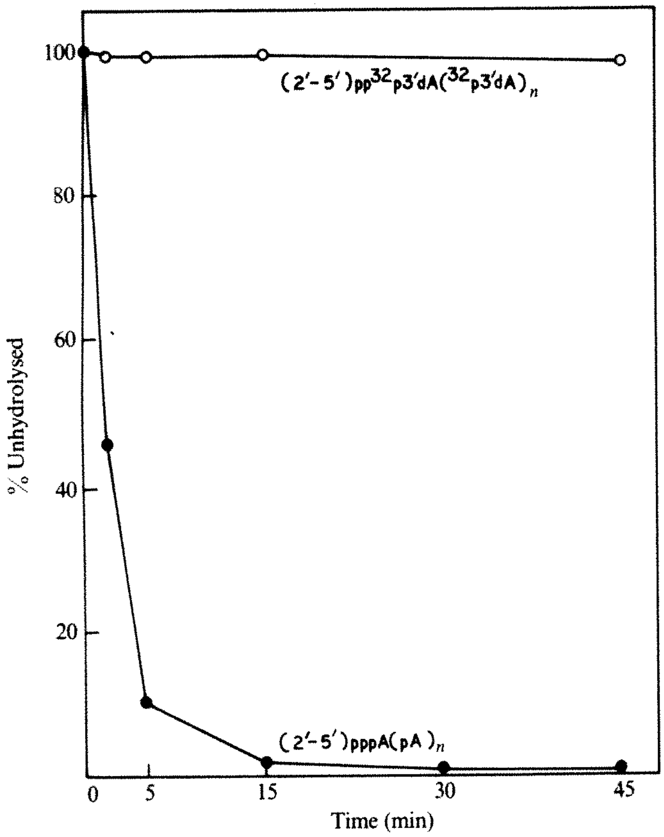


Fig. 3 Stability of (2'-5')pppA(pA)$_n$ and (2'-5')ppp3'dA(p3'dA)$_n$ in HeLa cell extracts. ^{3}H-(2'-5')pppA(pA)$_n$ and ^{32}P-(2'-5')ppp3'dA(p3'dA)$_n$ were synthesized as described in Table 1 legend. The oligonucleotides were obtained from the 0.35 M KCl eluate (DEAE-cellulose chromatography), desalted (Sephadex G-10) and concentrated *in vacuo*. The methods used to determine degradation of oligonucleotides were a modification of that of Minks *et al.*[13]. Incubation mixtures (0.125 ml) contained 0.6 parts of HeLa cell extract prepared as described previously[25], 2.5 mM Mg(OAc)$_2$, either ^{32}P-(2'-5')ppp3'dA(p3'dA)$_n$ (12 μM) or ^{3}H-(2'-5')pppA(pA)$_n$ (5 μM), 120 mM KCl, 20mM HEPES pH 7.4, and 1 mM dithiothreitol. Incubations were performed at 30°C. At the times indicated, 25-μl samples were withdrawn, heated for 3 min at 95°C, and the amount of undegraded oligonucleotide was determined by DEAE-cellulose chromatography as described previously[13]. Degradation of oligonucleotides was monitored as a decrease of radioactive material eluting with 0.35 M KCl compared with the zero time point (100% undegraded).

is important but not essential for the synthesis of (2'-5')pppA(pA)$_n$, (3) (2'-5')ppp3'dA(p3'dA)$_2$ is a more potent inhibitor of translation than (2'-5')pppA(pA)$_2$, (4) (2'-5')ppp3'dA(p3'dA)$_2$ and (2'-5')pppA(pA)$_3$ are approximately equal in their abilities to inhibit translation, and (5) enzymatically synthesized analogue (2'-5')ppp3'dA(p3'dA)$_n$ and chemically synthesized core analogue (2'-5')3'dA-(p3'dA)$_2$ are not hydrolysed by the 2', 5'-phosphodiesterase in HeLa cell extracts. 3'dATP is also converted to (2'-5')ppp3'dA-(p3'dA)$_n$ by extracts of human fibroblast interferon-treated L cells (manuscript in preparation). Enzymatically and chemically synthesized core (2'-5')3'dA(p3'dA)$_2$ inhibits transformation of Epstein–Barr virus (EBV)-infected lymphocytes in studies with human umbilical cord lymphocytes infected with EBV (manuscript in preparation).

The replacement of ATP with 3'dATP as a substrate for (2'-5')(A)$_n$ synthetase has been reported from three other laboratories. Lengyel and co-workers showed that (2'-5')(A)$_n$ synthetase from interferon-treated EAT cells can link adenylate moieties to the 2'-hydroxyl of 3'dATP[15]. Justesen *et al.* reported that 3'dATP is a chain terminator with respect to the addition of one 3'-deoxyadenylate moiety to either dimer or trimer (2'-5')pppA(pA)$_n$ following 3-h incubations with rabbit reticulocyte (2'-5')(A)$_n$ synthetase[16]. Minks *et al.* reported that 3'dATP inhibits the HeLa cell (2'-5')(A)$_n$ synthetase in the presence of ATP with respect to formation of (2'-5')pppA(pA)$_n$ (ref. 17). Baglioni and co-workers[18] have also reported that the ATP analogue, 1-N^6-ethenoATP, is converted to (2'-5')oligonucleotides which activate HeLa cell endonuclease and degrade vesicular stomatitis virus RNA. As (2'-5')oligonucleotides have been shown to inhibit mitogen-stimulated DNA synthesis in lymphocytes, as does interferon[19], it is tempting to speculate that analogues of (2'-5')pppA(pA)$_n$ may potentiate, extend or regulate these activities and could possibly be used as chemotherapeutic agents which might replace interferon.

The structure of the compound eluting at 29 ml on the DEAE cellulose column (Fig. 1*a*) is now under investigation.

We thank Nancy Lee Reichenbach for technical assistance in the isolation of cordycepin and the chemical synthesis of 3'dAMP and 3'dATP, Dr Wolfgang Pfleiderer for the supply of authentic chemically synthesized (2'-5')3'dA(p3'dA)$_2$,and Dr D. S. Rabson for human fibroblast interferon. This research was supported in part by NIH and NSF grants.

Received 10 April 1980; accepted 5 March 1981.

1. Hovanessian, A. G., Brown, R. E. & Kerr, I. M. *Nature* **268**, 537–539 (1977).
2. Kerr, I. M., Brown, R. E. & Hovanessian, A. G. *Nature* **268**, 540–542 (1977).
3. Kerr, I. M. & Brown, R. E. *Proc. natn. Acad. Sci. U.S.A.* **75**, 256–260 (1978).
4. Ball, L. A. & White, C. N. *Proc. natn. Acad. Sci. U.S.A.***75**, 1167–1171 (1978).
5. Schmidt, A. *et al. FEBS Lett.* **95**, 257–264 (1978).
6. Minks, M. A., Benvin, S., Maroney, P. A. & Baglioni, C. *J. biol. Chem.* **254**, 5058–5064 (1979).
7. Doetsch, P., Wu, J., Sheckman, G. D. & Suhadolnik, R. J. *Fedn Proc.* **39**, 1778 (1980).
8. Cory, J. G., Suhadolnik, R. J., Resnick, B. & Rich, M. A. *Biochim. biophys. Acta* **103**, 646–653 (1965).
9. Uematsu, T. & Suhadolnik, R. J. *J. Chromatogr.* **123**, 347–354 (1976).
10. Brownlee, G. G. in *Laboratory Techniques in Biochemistry and Molecular Biology* Vol. 3 (eds Work, T. S. & Work, E.) 1–265 (Elsevier, Amsterdam, 1972).
11. Charubala, R. & Pfleiderer, W. *Tetrahedron Lett.* **21**, 4077–4080 (1980).
12. Williams, B. R. G., Golgher, R. R., Brown, R. E., Gilbert, C. S. & Kerr, I. M. *Nature* **282**, 582–586 (1979).
13. Minks, M. A., Benvin, S., Maroney, P. A. & Baglioni, C. *Nucleic Acids Res.* **6**, 767–780 (1979).
14. Schmidt, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **76**, 4788–4792 (1979).
15. Samanta, H., Dougherty, J. P. & Lengyel, P. *J. biol. Chem.* **255**, 9807–9813 (1980).
16. Justesen, J., Ferbus, D. & Thang, M. N. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4618–4622 (1980).
17. Minks, M. A., Benvin, S. & Baglioni, C. *J. biol. Chem.* **255**, 5031–5035 (1980).
18. Baglioni, C., Maroney, P. A. & West, D. K. *Biochemistry* **18**, 1765–1770 (1979).
19. Kimchi, A., Shure, M. & Revel, M. *Nature* **282**, 849–851 (1979).
20. Suhadolnik, R. J. *Nucleoside Antibotics* 52 (Wiley, New York, 1970).
21. Suhadolnik, R. J. *et al. J. biol. Chem.* **252**, 4125–4133 (1977).
22. Fraenkel-Conrat, H. & Steinschneider, A. *Meth. Enzym.* **B12**, 243–246 (1967).
23. Martin, E. M., Birdsall, N. J. M., Brown, R. E. & Kerr, I. M. *Eur. J. Biochem.* **95**, 295–307 (1979).
24. Wu, J. M., Cheung, C. P., Bruzel, A. R. & Suhadolnik, R. J. *Biochem. biophys. Res. Commun.* **86**, 648–653 (1979).
25. Weber, L. A., Feman, E. R. & Baglioni, C. *Biochemistry* **14**, 5315–5321 (1975).

Erratum

In the letter 'Scale of body pattern adjusts to available cell number in amphibian embryos' by Jonathan Cooke, *Nature* **290**, 775–778 (1981), brackets in Table 1 were meant to refer to duplications of control or experimental embryos within each matched set. As these brackets are partially misaligned, they should be ignored and the data read by the use of a horizontal rule to lead the eye across the page in the absence of gaps between the data of each set.

Nature Indexes

Author indexes to *Nature* are published bimonthly, in the third week following the end of each volume. A complete subject and author index is published at the end of each year. Author indexes published this year are those for volumes 289 (26 march) and 290 (21 May).

MATTERS ARISING

Polygon patterns on Europa

PIERI[1], in his study of the lineament and polygon patterns on Europa, has discussed the problem of the development of crack networks in cooling basalt flows and similar planar, contracting systems. He observed that although an ideal hexagonal pattern maximizes the reduction of strain energy per unit surface area and is therefore the most efficient pattern given an isotropic stress field, nevertheless, even well developed columnar basalts tend to show many pentagons along with the expected hexagons.

The observations by Beard[2] on basalt flows the data used by Pieri) certainly show a large proportion of pentagons, ranging from 50% at Mount Rodeix to 35% at the Giants Causeway. Beard also recorded the mean number of sides of the basalt polygons and obtained values ranging from 5.23 at Mount Rodeix to 5.66 at the Giants Causeway. In an ideal system, with only 3-rayed crack vertices the mean side number should be six. A simple Monte Carlo model of a contracting cooling basalt can explain the significant occurrence of non-hexagons and also the reason why Beard's side counts were less than the ideal six.

In natural conditions, in an essentially unbounded basalt flow, one would expect the crack pattern to begin forming in a random manner; conditions can never be ideal enough for the entire pattern to develop in one cracking operation. Thus it can be expected to grow from randomly distributed stress centres[3]. Using a simple random number placement method and an effectively infinite field crack pattern, models can be produced which show many of the features observed by Beard. The model networks have a significant number of pentagons and heptagons and with just the random condition operating always have a mean polygon side number of six. However, a small number of extremely short sides are always produced and if these are eliminated the mean side value drops to ~5.6–5.7.

This suggests that Beard was unable to detect the very shortest sides and that this accounts for his mean side values in the 5.6 region. If this is so then a flow like the Giants Causeway is very close to an ideal flow; the ideal flow has a mean side number of six but as nature favours random processes we should expect a reasonable proportion of pentagons and heptagons. Pieri stated that the more isotropic and uniform the stress, the more likely it will be that hexagons form: this is looking to an ideal–ideal situation; in a real–ideal situation (a possible geological happening) it is likely that the mean side number will approach six but we should still expect non-hexagons to appear in the network.

I. J. SMALLEY

DSIR Soil Bureau,
Lower Hutt,
New Zealand

1. Pieri, D. C. *Nature* **289**, 17–21 (1981).
2. Beard, C. N. *Bull. geol. Soc. Am.* **70**, 379–381 (1959).
3. Smalley, I. J. *Geol. Mag.* **103**, 110–114 (1966).

PIERI replies—It is clear from Smalley's comment above and from previous work[1] that polygonal fracture patterns forming in real materials in even the most ideal of isotropic stress regimes are subject to compositional anisotropies at many scales and a fracture network cannot form all at once. Thus, the ideal perfectly-formed pure-hexagonal pattern will never form in nature and will have particular difficulty in the context of global planetary fracture patterns[2]. I did not mean to imply that the end result of optimized real conditions would be pure hexagonal patterns, but rather that hexagons would be more numerous in situations where conditions were more uniform. Clearly, pentagons and heptagons will still be present, and I thank Smalley for clarifying this point.

The research described in this note was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under contract with NASA.

D. C. PIERI

Jet Propulsion Laboratory,
California Institute of
Technology,
Pasadena, California 91109, USA

1. Smalley, I. J. *Geol. Mag.* **10**, 110–114 (1966).
2. Pieri, D. C. *Nature* **289**, 17–21 (1981).

Low-angle X-ray scattering of chromatin

LANGMORE AND SCHUTT[1] have recently reported low-angle X-ray scattering experiments on chicken erythrocytes in which they find a peak at ~400 Å in the plot of s^2I against s. They conclude that this peak is related to the side-by-side packing of chromosome fibres. I wish to show here that this peak may have a different interpretation.

Langmore and Schutt[1] multiply the recorded intensity I by s^2 to correct for the random orientation of the fibres with respect to the X-ray beam, where $s = 2\sin(\theta/2)/\lambda$. However, as shown quite generally by Porod[2] and discussed in detail by Mittelbach[3], information on the cross-section properties of rod-like particles is best obtained from the study of the dependence of Is (instead of Is^2) on s.

At very low angles, a system of rod-like particles obeys the approximate Guinier equation:

$$(Is) = K\exp[-2(\pi s R_G)^2] \qquad (1)$$

where K is a constant and R_G the radius of gyration of the cross-section. For a cylindrical particle of uniform density and radius R, $R_G = 0.707\,R$. If Is^2 is plotted instead of Is, a maximum in the curve will appear when $d(Is^2)/ds = 0$. Introducing this condition in equation (1), it is easily shown that a maximum in the plot of Is^2 against s should appear for

$$s = 1/2\pi R_G \qquad (2)$$

Therefore, the maximum found by Langmore and Schutt[1] in the plots of Is^2 against s for $s = 1/400\ \text{Å}^{-1}$ is consistent with the presence of cylindrical objects with a cross-sectional radius of gyration $R_G = 64$ Å, as calculated from equation (2). This value would correspond to cylindrical particles of 180-Å diameter if their electron density were uniform. Thus I think that the peak observed by Langmore and Schutt[1] is due to intrinsic features of the chromatin fibres and does not demonstrate any regular side-by-side packing of chromosome fibres.

I suggest that the determination of the position of maxima and the use of equation (2) may be a useful alternative to Guinier plots in determining the value of R_G. A similar procedure could be used for globular particles, but in that case the maxima in the plot of Is against s would be used.

JUAN A. SUBIRANA

Unidad de Quimica Macromolecular,
Consejo Superior de Investigaciones
Cientificas,
Escuela Técnica Superior de
Ingenieros Industriales,
Diagonal 999, Barcelona, Spain

1. Langmore, J. P. & Schutt, C. *Nature* **288**, 620–622 (1980).
2. Porod, G. *Acta phys. austriaca* **2**, 255–292 (1948).
3. Mittelbach, P. *Acta phys. austriaca* **19**, 53–102 (1964).

LANGMORE AND SCHUTT REPLY—Although it is necessary that widely spaced homogeneous cylindrical fibres of diameter D give rise to a maximum in s^2I at $\sim s = (2.2\,D\ \text{Å})^{-1}$, a peak in the s^2I curve is not sufficient to prove the existence of widely spaced homogeneous cylindrical particles. The analysis that Subirana suggested is only relevant to the determination of the radius of gyration of particles provided that these particles are rod-like and widely separated. Proper Guinier analysis[1] tests the validity of these

0028-0836/81/220359—02$01.00

necessary conditions by comparing the shape of the experimental scattering curve with that predicted for widely separated rod-like particles. Such analysis of the scattering from our molecules shows that the Guinier conditions are not met, because the experimental plot of log sI against s^2 is not linear, due to the fact that the chromosome fibres we have studied are not widely separated (*vide infra*).

In our letter we outlined three basic arguments to prove that the 400 Å band is due to the packing of fibres and therefore not the result of scattering from widely separated rod-like molecules. These arguments were that: (1) the 400-Å peak is found in the I versus s plots of living chicken erythrocytes after subtraction of the non-nuclear background (represented by the scattering from living rabbit erythrocytes); (2) the 400-Å band in the s^2I vs s plots can be directly related to the 300-Å centre-to-centre spacing of fibres often seen in thin sections of a wide range of cells, including chicken erythrocytes[2]; and (3) the 400-Å peak in the s^2I curves of the scattering from intact nuclei is lost when the thick fibres are intentionally disaggregated into soluble thick fibres (as assayed by our electron microscopy) by elimination of divalent cations from the buffer. Each of these arguments disproves that scattering from isolated thick fibres has given rise to the 400-Å features in our patterns.

If, as we suggest, erythrocyte chromosome fibres are homogeneous cylinders of 370–400 Å diameter, a dilute dispersion of fibres would give rise to a maximum in s^2I at ~850 Å. To see a peak at these very small angles would require better first-order resolution in our X-ray camera.

Thus we feel that our experiments have demonstrated that the 400-Å feature is due exclusively to the side-by-side packing of chromosome fibres. The absence of a 400-Å feature in disaggregated fibres is inconsistent with the hypothesis put forth by Subirana. Furthermore the analytical technique he has proposed is not generally valid because it cannot test whether the necessary conditions for a Guinier analysis are met.

JOHN P. LANGMORE

Division of Biological Sciences
and Biophysics Research Division,
University of Michigan,
Ann Arbor, Michigan 48109, USA

C. SCHUTT

MRC Laboratory of Molecular
Biology,
Hills Road,
Cambridge CB2 2QH, UK

1. Guinier, A. & Fournet, G. *Small Angle X-Ray Scattering* (Wiley, New York, 1955).
2. Davies, H. G. *J. Cell. Sci.* **3,** 129–150 (1968).

Stability of loess in light of the inactive particle theory

A RECENT article by Smalley *et al.*[1] has suggested an interesting theory pertaining to loess stability. The inactive particle theory of soil sensitivity attempts to explain the sensitivity of quickclays on the basis that these soils possess only small amounts of clay minerals and as a result, develop no long-range bonds. Therefore once disturbed, they do not retain the adequate shear strength needed to remain stable[2,3].

In the classical Terzaghi sense, loess soils, which are composed predominantly of silt particles, would probably not be considered sensitive soils, in that the ratio of undisturbed to remolded strength, at constant moisture content, is usually around 3, depending on clay content, and therefore would generally fall into the category of medium sensitivity. If, however, sensitivity is taken as the ratio of undisturbed to saturated strength (in unconfined compression) as indirectly suggested by Feda[4], then some loess soils would have to be considered quick.

Loess soils, particularly bluff-line deposits, typically lose shear strength as a result of moisture saturation; in which case landslide potential becomes an extreme hazard. Loess in this state may be susceptible to "spontaneous liquifaction"[5] which could help explain the extent of landslides during the 1920 earthquake in the Kansu Province of China.

In recent investigations throughout the midwestern US, *in-situ* stability of loess has been related to liquidity index, that is when the *in-situ* moisture content reaches the liquid limit, usually on saturation, stability is all but lost and can be readily identified by isolated flow in boreholes. Because liquid limit is related to clay content, and saturation moisture is a function of density, this instability only occurs in special circumstances, typically low density and low clay content. As both calcareous and leached examples of this condition have been noted, the leaching theory of Rosenquist does not seem applicable in that carbonate leaching need not be complete, however, the cementation bond may be weakened. Scanning electron microscopy (SEM) qualitatively reveals that only a small portion of the clay fraction (<2 μm) of loess is composed of clay minerals, with the remainder being comprised of clay sized quartz particles. The amount of clay is related to closeness of the deposit to the source. An indirect measure of sensitivity is given by Skempton's[6] "activity" which for bluff-line loess averages ~0.46, clearly in the "inactive" range.

The inactive particle theory seems to have significance to loess stability, however, to investigate the theory fully thermogravimetry will be necessary to quantify the amount of active minerals present in each of the various particle-size fractions.

A. J. LUTENEGGER

Geotechnical Research
Laboratory,
Iowa State University,
Ames, Iowa 50011, USA

1. Smalley, I. J., Ross, C. W. & Whitton, J. S. *Nature* **288,** 576–577 (1980).
2. Smalley, I. J. *Nature* **231,** 310 (1971).
3. Cabrera, J. G. & Smalley, I. J. *Eng. Geol.* **7,** 115–133 (1973).
4. Feda, *J. Eng. Geol.* **1,** 201–219 (1966).
5. Terzaghi, K. *Applications of Geology to Engineering Practice*, (Geological Society of America, 1950) (ed. Paige, S.) 83–123.
6. Skempton, A. W. *Proc. 3rd Int. Conf. Soil Mech. Foundn Eng.* Vol. 1, 57–61 (1953).

SMALLEY REPLIES—LUTENEGGER makes quite a good case for considering loess a sensitive soil, and for applying the inactive-particle short-range-bond hypothesis to it, but if we are going to take this approach we shall have to acknowledge the pioneering opinions of Denisov[1]. He stated that a subsident loess (the sort described by Lutenegger) is extremely sensitive, particularly when saturated. The water content in such a state is above the liquid limit; the condition which occurs in the classic quickclays. Denisov pointed out that loess in this state preserves some strength and can form steep sides to canals and pits. When disturbed the shear strength drops to zero. The basic reason, according to Denisov, for the appearance of highly sensitive quickclays is their "preservation under natural conditions of the uncompressed state": an observation which could be very close to the truth.

I. J. SMALLEY

New Zealand Soil Bureau,
Lower Hutt, New Zealand

1. Denisov, N. Ya. *Osnov. Fundam. Mekh. Grunt.* **5,** 5–8 (1963).

Matters Arising

Matters Arising is meant as a vehicle for comment and discussion about papers that appear in *Nature*. The originator of a Matters Arising contribution should initially send his manuscript to the author of the original paper and both parties should, wherever possible, agree on what is to be submitted. Neither contribution nor reply (if one is necessary) should be longer than 300 words and the briefest of replies, to the effect that a point is taken, should be considered.

0028–0836/81/220360—02$01.00

BOOK REVIEWS

Sowing the seeds of an empire

A. Hunter Dupree

Sir Joseph Banks: 18th Century Explorer, Botanist and Entrepreneur. By Charles Lyte. ISBN 0-7153-7884-8. Pp.248. (David and Charles: 1980.) £10.50, $32. *The Journal of Joseph Banks in the Endeavour 1768–1771.* With commentary by A.M. Lysaght. Facsimile edition, 2 vols, limited to 500 copies. Vol.I, pp.616, ISBN 0-904351-03-3; Vol. II, pp.732, ISBN 0-904351-04-1. (Genesis Publications, Guildford, Surrey: 1980.) £230.

FOR a century, from 1760 to 1860, the dominant pattern for a scientific career little resembled the academic progression of the twentieth century. A taste for genteel collecting and an irrelevant university or almost irrelevant medical education was the total preparation for youths of scarcely twenty to embark on an epic journey or ocean voyage. By the time they were homeward bound, the still-young men were not only rich in experience but splendidly and broadly educated scientists, capable of thinking in global terms. Typically, an intense period of working on the collections of the voyage led to a geographically static later life. Darwin's career is an extreme case of this pattern, since he never left Britain after the return of the *Beagle* in 1836 at the age of twenty-seven. A decade later, T.H. Huxley used the voyage of the *Rattlesnake* for his scientific and worldly education. The prototype of this career pattern was Joseph Banks's circumnavigation of the world betwen 1768 and 1771, as the botanist on that prototype of the scientific exploring expedition, Captain James Cook's first voyage to the Pacific.

Banks, descended from a well-to-do Lincolnshire family, was an indifferent scholar at Harrow, Eton and Christchurch, Oxford, but his interest in natural history led him in 1766, when he was only twenty-three, to go to Newfoundland and Labrador aboard a naval vessel. A.M. Lysaght's *Joseph Banks in Newfoundland and Labrador, 1766* (Faber & Faber: 1971), is a beautiful and definitive book, with an analysis of the natural history and a collection of diaries and documents.

The young Banks paid his own way, and that of his illustrators and collectors, as botanist on Cook's expedition to observe the transit of Venus at Tahiti in 1769. A pattern of specialization, both among scientific disciplines and between naval officers and civilian scientists, was already well established. Cook the hydrographer and geographical explorer would have accomplished less without the civilians, and an astronomer alone could not have brought home the contributions to botany, zoology and anthropology which Banks's personal skill and financial resources made possible. Thanks to *The Endeavour Journal of Joseph Banks 1768–1771*, edited by J.C. Beaglehole (Angus & Robertson: 1962), Banks's account of the voyage has been available to scholars around the world, bringing to bear in extensive notes the scholarship of many experts, including Averil Lysaght.

Courtesy Mr David Hughes. Photograph by John Bignell.

Joseph Banks, detail of a portrait by Francis Cotes. The painting was only recently rediscovered and is reproduced for the first time in the facsimile edition of the Endeavour *journal. Executed in the middle 1760s, the portrait is of interest as the only pictorial record of Banks before he embarked in the* Endeavour.

Banks's enthusiasm for the role which he had assumed on Cook's first voyage made his plan for the second so expansive that he came into conflict with Cook and the Navy Board, and the sailing qualities of the ship he sought to modify added further problems. When the superiority of military command to scientific command was decisively confirmed, Banks and his entourage stayed at home.

He assuaged his injured pride by a brief private expedition to Iceland in 1772, returning to England to remain almost without exception until his death in 1820. A friend of King George III, he became adviser to the King on the development of the Royal Botanical Garden at Kew and informally its permanent director, and President of the Royal Society of London in 1779. In a review of Beaglehole's edition of the *Endeavour* journal in the *Times Literary Supplement* (July 20, 1962, p.523), the reviewer said:

> He [Banks] was to become a sort of scientific adviser at large, a central clearing house of other naturalists' collections and ideas, perhaps, indeed, the first of the scientific civil servants; and in any case, in his long reign as President of the Royal Society, the exact opposite of the perpetual global tourist for which he had first cast himself. These extended views of his long career need much more careful sorting out.

That call for a well-rounded biography doing justice to science policy was well taken in 1962 and remains unfulfilled today.

The emphasis of Lyte's new biography seriously detracts from its usefulness. Seventy per cent of the book brings one to November 1772, when Banks at the ripe age of twenty-nine returned from his Iceland expedition. Hence scholars can without loss continue to turn to Beaglehole and Lysaght, or even to the previous biographies by H.C. Cameron (1952) and Edward Smith (1911). Lyte does not delve below the surface in his analysis either of the natural history or the science policy that makes Banks such a significant figure.

In spite of this criticism of the scholarly contribution of Lyte's book, one can point out both that it reads pleasantly and that the selection of episodes for emphasis presents an implicit interpretation of some interest. Banks, by this reading, became Cook's negotiator with the peoples on the shore. His anthropological research had the very practical purpose of gaining supplies for the ship and access for landing. The shore parties could then proceed with science, trade and, if possible, sex. The sequence of negotiations often involved the use of firearms by the British, and the show of force frequently produced casualties. Banks then awaited a second approach from the natives which would allow him to move confidently among them, trading nails for supplies, becoming adept in the language and quickly identifying the important people with whom to deal. All of the negotiation was preliminary to vigorous collecting of cultural artifacts, animals and especially plants, including horticultural varieties and weeds — a valuable record for students of the worldwide diffusion of organisms associated with man. Hence natural history became fused with the national and military goals of the expedition, and Banks the botanist became second only to Cook in shaping the policy of European penetration into the southern Pacific — New Zealand and Australia as well as the Polynesian islands, of which Tahiti was the most important. An element in this policy was persuading a likely native or two to continue with the ship to serve as interpreter and guide. The links between scientific, military and national interests, represented by useful plants and animals, became a part of the system by which Britain replaced her lost American colonies with a world empire.

The Journal of Joseph Banks in the Endeavour, with a commentary by A.M. Lysaght, is a facsimile of the original manuscript in the Mitchell Library in Sydney, Australia. The commentary breaks new ground with reflections on Banks as an anthropologist. The extensive index takes full account of Banks's lapses in pagination and provides firm control of both scientific and geographical names. Only the most specialized of scholars will have occasion to consult in detail the facsimile itself; there is, however, a more important justification for this expensive, limited edition with its fine binding and gilded pages. Not only were the Banks manuscripts withdrawn from the British Museum in 1886 and scattered all over the world by an auction at Sotheby's. Banks himself placed a burden on posterity by allowing the illustrations of the *Endeavour* voyage by Sydney Parkinson and others to go unpublished, even though the copper plates and the descriptions by D.C. Solander were ready by the time of Solander's death in 1782. An important function of the rare book library is to bring these fugitives to light. An alternative way of financing the publication of residues of scientific literature left behind in enclaves for two centuries is hard to conceive, unless unusual beauty or interest attracts the bibliophile. These volumes fit neatly between *Captain Cook's Florilegium*, published by Lion and Unicorn in 1973, and the *Banks' Florilegium* recently announced by Alecto Historical Editions, which at long last reproduces in full colour the engravings of the drawings by Sydney Parkinson. The facsimile edition contains an extensive list of the illustrations. Although Banks the President of the Royal Society has yet to find his biographer, the Banks of Cook's first voyage is now increasingly well-served by the literature. □

A. Hunter Dupree is Professor in the Department of History at Brown University, Providence, Rhode Island.

Princeton's obeisance to Einstein

Paul Davies

Some Strangeness in the Proportion: A Centennial Symposium to Celebrate the Achievements of Albert Einstein. Edited by H. Woolf. Pp.538. ISBN 0-201-09924-1. (Addison-Wesley: 1980.) $43.50, £26.10.

EINSTEIN'S conception of the physical world was so broad, it is appropriate that a centennial symposium should deliberate on topics as diverse as quarks and quasars. The hundredth anniversary of the great man's birth was marked by a variety of international gatherings and the publication of several books. These activities ranged in subject across Einstein's political and religious convictions, his contributions to statistical mechanics, relativity and quantum theory, his philosophical deliberations and his cosmological speculations. Princeton's symposium had some of each.

Science in general, and physics in particular, is often presented as an impersonal, logical tide of progress, rooted in narrowly defined procedures of experimental test and mathematical precision. The reality is different. Like so many human activities, the development of physics revolves around personalities, with their attendant prejudices, beliefs and jealousies. Although ultimately resting on empirical foundations, the actual cut and thrust of mainstream research is dictated by individuals and mirrors thus their tastes and aspirations. This individualism is reflected in the scientific community's creation of idols. Great physicists — Newton, Laplace, Maxwell, Fermi, etc., etc. — are exalted, almost deified by succeeding generations of their profession. Einstein is one of the few scientists whose awesome genius attracted the fascinated adulation of a much wider public: the Gandhi or Napoleon of physics.

The impact of Einstein's theoretical work on the subsequent development of physics is legendary. At a stroke he helped initiate two revolutionary new conceptual structures — special relativity and the quantum theory — and completed one of them (his general theory of relativity) so decisively that he left his colleagues with little more fundamental work to do on the subject for the next 40 years. Because these theories, mostly presented between 1905 and 1915, later assumed such immense importance for physics, almost every facet of modern research incorporates Einstein's ideas somehow. Nevertheless, science moves on. One detected, perhaps, a somewhat strained quality in the universal claims of the participants that their particular branch of galactic dynamics or gauge theory or particle physics is but a natural and inevitable extension of the patriarch's great intellectual empire. What would Einstein really have made of gluons or gravitinos?

How does one celebrate Einstein's achievements? Scientifically the formula is universal: astrophysics and particle physics, with special emphasis on their intersection. Undoubtedly the event of the centennial epoch was the discovery of gravitational radiation, predicted by Einstein himself over 60 years ago, and persuasively confirmed by the recent observations of the so-called binary pulsar. On the theoretical front the swift progress towards unifying the four fundamental forces of nature dominates the scene. At the Princeton symposium the gravitational component of this confluence — supergravity — provided a natural complement to the many experimental papers.

In spite of these specialist details, the perspective was all-embracing. "The Universe" was reviewed and analysed, its mysteries and paradoxes set alongside its simplicity and splendour. Where did the cosmic material come from, where is the antimatter, why is the expansion so uniform, what happened in the beginning, how will it end . . .? The contributors were all experts of international renown, and between them they covered all the puzzles, problems, successes and failures of modern cosmology.

Special relativity was not submerged by the general theory. Wolfgang Panofsky explained how the principles of the special theory are these days a matter for routine engineering work, in such devices as high-energy klystrons and accelerators. I suspected a side-swipe at the malcontents who still insist, all these years on, that there is something fishy about time dilation, length contraction and all the other weird effects predicted by Einstein's theory. Most memorable of Panofsky's comments is that to a high-speed electron, Stanford's two-mile linear accelerator seems only about 30 inches from end to end!

For many readers the more attractive part of these symposium proceedings will relate to the papers about Einstein himself, and those heady days in the first quarter of this century when science was turned upside down and major advances jostled each other in the literature. For today's physicists, groping among the complexities of a tangled and increasingly fragmented subject, the anecdotal images of a serene intellect nimbly picking its way through half-a-dozen major scientific conundrums makes compelling reading. Of course, he was not always right. His implacable opposition to the claims of Bohr, Born and others that quantum theory was a complete description of nature made him a curiously isolated figure — "like an ostrich", he himself remarked, unable to face the "evil quanta".

"I knew Einstein", in spite of its I-danced-with-the-Prince-of-Wales overtones, is surely an impressive and riveting revelation. Several of the symposium participants either knew, or at least spoke to this Swiss–German–American Jew. No narrow-minded scientist he, we are assured. Einstein liked (even played!) chamber music, discussed Spinoza, rode on trams. He also lived through a time of intense political and social upheaval, and championed a variety of causes.

Remembering Einstein evokes among physicists a curious amalgam of humility and conceit. Who could ever aspire to such a complete mastery over their chosen subject? Measured against the Einstein standard, we are all lesser scientists. Yet we, the community of physicists, claim Einstein as one of our own, and love to dazzle each other, and the wider public, with the brilliant advances that Einstein stimulated.

Was Einstein a true genius who would have succeeded in any chosen activity, or was he merely a man of the moment, thinking in the right way at the right time about the right sort of problem? If he lived today, would his contributions to physics be as profound and far-reaching? Somehow I do not think so. Yet that is not to demean his achievements, which are paralleled only by Newton. Because of Einstein the world is a far richer place.

This centennial volume, like the others, is a patchwork. The joins are smoothed somewhat by augmenting the technical papers with comments from subsidiary speakers, together with a few rather unsatisfactory open discussions involving a good deal of rambling. It is, however, a delightful addition to one's library, and an invaluable source of Einstein stories. □

Paul Davies is Professor of Theoretical Physics at the University of Newcastle upon Tyne. His latest book is The Edge of Infinity *(Dent, 1981).*

Limb diversity, what can be more curious?

Julian Lewis

The Development of the Vertebrate Limb: An Approach through Experiment, Genetics, and Evolution. By J.R. Hinchliffe and D.R. Johnson. Pp.266. ISBN 0-1985-7552-1. (Clarendon/Oxford University Press: 1980.) £20, $59.

THE question of the relationship between evolution and embryonic development is forever tantalizing. The experts on each of these topics feel that somehow they hold the key to the understanding of the other; and yet it is remarkably hard to translate the feeling into a real and useful new insight. Hinchliffe and Johnson preface their monograph with an apt quotation from Darwin:

> What can be more curious than that the hand of a man, formed for grasping, that of a mole for digging, the leg of a horse, the paddle of the porpoise, and the wing of the bat, should all be constructed on the same pattern, and should include similar bones, in the same relative positions?

Though limb proportions have varied, limb topology has remained practically constant. This, as Darwin saw, is evidence that structures evolve in a quasi-continuous fashion, by a succession of small adjustments, rather than by big, single leaps that change the qualitative plan. The important implication for the developmental biologist is that even though the quantitative parameters of limb development differ from species to species, the qualitative principles can be expected to remain almost the same. The search for general mechanisms rests on that proposition. But it is still a proposition that needs to be substantiated directly. Thus Hinchliffe and Johnson's book, though it does not offer any new theoretical insights, is valuable for the comparative information that it provides.

The authors begin with an account of the evolution and of the diverse adaptations of the vertebrate limb. The core of the book is a review of descriptive, comparative and experimental studies of limb development and regeneration. The final chapter discusses limb mutations. The book is well-written, informative and at times even entertaining. It is, however, rather uneven in its coverage, and, like many reviews, it suffers from the lack of a unifying central thesis. I felt this lack particularly in the chapters on development. These focus on the genesis of the limb skeleton, and touch on many other interesting topics, including some fundamental problems of developmental biology; but the account does not seem to be informed by any clear general conception of the principles of development. Perhaps the authors simply wish to be objective, and to keep their prejudices from obtruding. But it is consquently hard to tell which experiments are to be considered crucial, which models important, which theories right. To borrow a phrase from Hinchliffe and Johnson, I fear that the general reader may be left asking: "Amidst this welter of morphological anomaly and theoretical speculation, what conclusions can we draw . . .?".

The declared aim of the authors is to provide "a reference book which is at the same time an account of the current 'state of the art'." As such the book should be useful; the experts will be able to bring their own prejudices to the reading of it.

Julian Lewis is a Lecturer in the Department of Anatomy at King's College, University of London.

BIOCHEMICAL JOURNALS FROM GORDON AND BREACH:

Cancer Biochemistry Biophysics
Editor: Harry D. Brown,
Rutgers University, USA
This unique journal serves as a rapid communications medium for biochemists and biophysicists working in areas germane to cancer research. Its contents, international in scope, reflect the increasing impact of cancer investigation on other problems in biochemistry and biophysics.
Subscription rates
£78/$187.50 per vol. of 4 issues
Sample copies available on request.

ANNOUNCING A NEW JOURNAL...

Chemistry in Ecology
Editor: E. J. Perkins,
Marine Laboratory, University of Strathclyde, Scotland
The role of chemistry in ecological processes is the broad concern of this forthcoming journal, which will also examine biological interactions, waste disposal (including toxicology), and the ways in which the health of species and communities are thereby affected.
First issue to be published in August 1981.
Subscription rates
£83/$199 per vol. of 4 issues

Toxicological and Environmental Chemistry
Editors: O. Hutzinger, University of Amsterdam, and R. W. Frei, The Free University, Amsterdam
This interdisciplinary journal considers the biological and environmental effects of pesticides; air and water pollutants; toxic natural products; radioactive materials; and carcinogenic and mutagenic chemicals which display long-term biological effects.
Subscription rates
£83/$199 per vol. of 4 issues

The sterling prices are absolute.
The dollar prices are subject to exchange rate fluctuations. For our forthcoming JOURNALS CATALOGUE, please write to our Information Office

GB **Gordon and Breach Science Publishers**
42 William IV Street
London WC2N 4DE, UK

Circle No.36 on Reader Enquiry Card.

The promise and the problems of enzyme technology

Peter Dunnill

Immobilized Enzymes in Analytical and Clinical Chemistry: Fundamentals and Applications. Chemical Analysis, Vol.56. By P.W. Carr and L.D. Bowers. Pp.460. ISBN 0-471-04919-0. (Wiley: 1980.) £31.10, $63.20. *Immobilized Enzymes: An Introduction and Applications in Biotechnology.* By M.D. Trevan. Pp.138. ISBN 0-471-27826-2. (Wiley: 1980.) £8.75, $24.10.

THE great development which the field of immobilized enzyme technology has undergone is reflected in quite different ways in these two books. Ten years ago the field was so specialized that there was little need for a simple introductory text, but Michael Trevan's slim volume is published at a time when many people, including teachers, industrialists and even investors, need to know about the implications of enzyme technology, and have some grasp of the underlying principles. Equally, a subject such as the use of immobilized enzymes in analysis, which a decade ago warranted only a review chapter, now not only justifies a book, but Professors Carr and Bowers have had to add supplementary reference lists to deal with the outpouring of recent papers.

Dr Trevan divides his introductory account into three main parts. The first, dealing with the techniques of immobilization and its consequences for enzyme activity, is noteworthy for its clear and readable explanation of immobilized enzyme kinetics and the consequences of diffusional effects. The second reviews the current and potential applications of immobilized enzymes. This survey is comprehensive, but even more lead references would help to avoid any frustration in readers who wish to pursue more detailed information. The final section reflects renewed interest in the relationship between artificially immobilized enzymes and the functioning of enzymes which are fixed spatially within the cell, for example by association with membranes. Though the models which can be constructed to mimic active transport and other cellular processes are as yet crude, they point the way to exciting scientific developments. Introducing the possible industrial consequences of multi-enzyme catalysis to biochemistry students at the earliest possible stage could help to knit further the scientific and technological aspects of the subject.

The monograph by Carr and Bowers is a contribution to a series on analytical chemistry and sets out to provide a self-contained explanation of the use of immobilized enzymes in analytical and clinical chemistry. Therefore it deals with basic enzymology and enzyme kinetics, and with appropriate immobilization techniques. Like Trevan's book, a substantial part of the volume has to be devoted to the kinetics of immobilized enzymes, since a grasp of this aspect is central to interpretation of analytical results; however, as one would expect, the treatment is of greater depth and rigour, and this extends to the discussion of individual analytical methods where the focus is rightly on the underlying principles. The authors accept that immobilized systems have not made as much headway as expected, given the potential benefits in terms of reduced costs and increased accuracy. They ascribe this in part to the additional complexity of the two-phase enzyme–reagent system and to the need for a new method to show substantial advantages in order to displace established practice: anyone who has observed the pressures on hospital analytical facilities will know that introducing untried innovations is unlikely to be popular. However, the book's discussion of the strengths and weaknesses of immobilized enzymes for analysis should provide a firm foundation for the development of commercial systems.

Both books succeed in achieving their very different objectives and are valuable additions to the growing library of enzyme technology literature. □

Peter Dunnill is Reader in Biochemical Engineering at University College, University of London.

Pollution assessment: objectives and means

F. Moriarty

Biological Monitoring for Environmental Effects. Edited by Douglas L. Worf. Pp.225. ISBN 0-669-03306-5. (Lexington Books: 1980.) £15.50, $22.95.

THIS is the report of a conference and workshop held in North Carolina early in 1978. It was an all-American affair, and there is a lot of detail in the 22 papers that will not be of much interest to readers from other countries. Nevertheless, this book does raise several themes of general relevance for those who are concerned with monitoring the environmental effects of pollutants.

Biological monitoring was taken to mean the use of biota both to monitor the structure and function of ecosystems, and to measure the amount and distribution of contaminants. Several participants were concerned that current and impending regulations rely solely on chemical and physical measurements: the possibilities of, and need for, biological monitoring are not appreciated. Cairns and Dickson both emphasized that control of pollution by the "best applicable technology" is inappropriate, because it ignores the environment's capacity to deal with waste materials, and can result, in different circumstances, in either too much or too little protection of the environment.

A linked theme is the question of objectives. An adequate monitoring scheme cannot be designed until precise objectives have been stated. Weber discusses one aim of the Federal Water Pollution Control Act, to restore and maintain the biological integrity of the nation's waters. Biological integrity is not explicitly defined, and unless it means that ecosystems should remain in, or be restored to, their pristine condition, it is a nebulous concept that does not suggest objectives for monitoring. In practice, it is taken to mean fish, shellfish, wildlife and the effects of pollutants on the diversity, productivity and stability of communities of indigenous aquatic organisms. Osburn refers to the National Environmental Policy Act, which clearly accepts values such as aesthetic and cultural pleasure, and which can be translated into clear objectives.

Hirsch emphasizes that human impact on wildlife is synonymous with effects on ecosystems, discusses the scientific problems and then lists three institutional difficulties: lack of assured long-term funds for relevant ecological research, inadequate communication between administrators and ecologists, and lack of a focal point for expertise on ecological problems.

Other papers describe a range of techniques for monitoring in aquatic and terrestrial habitats, and the last five summarize the conclusions of five workshops. Current research is criticized for producing too many data of uncertain relevance; in addition more taxonomic skills are needed and a better ability to communicate scientific data to administrators. Research is also hampered by academic and administrative arrangements that militate against the development of multidisciplinary teams.

Several recent books have dealt with various aspects of monitoring. The range of topics in this book is limited, and much of the detail is of relatively parochial interest, but the problems encountered have a much wider relevance.

F. Moriarty is a Principal Scientific Officer at Monks Wood Experimental Station, Institute of Terrestrial Ecology, Abbots Ripton, Huntingdonshire.

Provocative images through the microscope

P. W. Hawkes

Scanned Image Microscopy. Edited by Eric A. Ash. Pp.461. ISBN 0-12-065180-7. (Academic: 1981.) £18.40, $49.

So accustomed are we to all the conveniences of two-dimensional image formation at optical wavelengths — the existence of lenses and of high-resolution polychromatic recording media in particular — that it is easy to forget that information can often be usefully conveyed by other fields for which no such facilities exist. The absence of lenses for X-rays, at least until very recently, is an obvious example. In such cases, the information may well be useful enough for other means of collecting it to be developed, and this is one of the reasons why the scanning principle is used. Its most obvious advantage arises when the incident radiation can be concentrated into a small spot on the specimen but cannot be focused after interacting with the latter, so that there can be no question of conjugate planes: acoustic microscopy is an example of this. Here, the specimen is scanned mechanically under the spot and a signal collected for each position.

Scanning is thus used for similar reasons in quite different domains, and this volume of Rank Fund Conference Proceedings contains accounts of four types of scanning microscopy: acoustic, optical, photoacoustic and soft X-ray microscopy. Scanning electron microscopy is excluded, on the grounds that it is now too well-known, which is true of the traditional reflection instrument but not of its much newer and still rather uncommon transmission counterpart. It is a pity that some contributions on the latter were not solicited, for the resemblances and differences between the scanning transmission electron microscope and the scanning optical instrument are interesting and this would have been an excellent opportunity to compare and contrast.

Even with this minor cavil, however, the coverage is thorough. Each of the four main sections contains chapters by the inventors or leading exponents of the instrument in question. Thus acoustic microscopy is discussed by C. F. Quate, H. K. Wickramasinghe and B. Nongaillard, as well as by several other contributors concerned with specific applications. Scanning optical microscopy is examined by W. T. Welford, G. J. Brakenhoff and colleagues, and by T. Wilson and C. J. R. Sheppard, among others. The other two sections have no less distinguished a cast. The result is that although the book has all the faults inherent in such proceedings' volumes — much repetition in the chapter openings in particular — it does provide a clear and readable introduction to these four scanning-based types of microscopy and gives a good idea of their potential. I suspect, too, that practitioners of each technique will find it an extremely useful guide to what is happening in related fields, particularly when these are used to complement one another (the example of scanning optical and electron microscopy, cited in the optical section, is one of many).

In short, this is an extremely useful collection: the individual contributions are long enough for the newcomer to be able to follow and the choice of scanning as leitmotiv makes good physical sense. Apart from the omission mentioned above, the absence of one other topic is to be regretted: spectroscopy using orthogonal transforms. Here, it is the need to "scan" a field with small detectors that has provided the incentive to seek new solutions, resulting in Hadamard transform techniques for example. Nevertheless, it is probably because the material that E. A. Ash has assembled for us is up to date, well presented and frequently provocative that we should have liked a little more. ☐

P. W. Hawkes is Maître de Recherches at the Laboratoire d'Optique Electronique du CNRS, Toulouse.

Pelagic predators

Geoffrey Fryer

Predation and Freshwater Communities. By Thomas M. Zaret. Pp.187. ISBN 0-300-02349-9. (Yale University Press: 1981.) £9.50, $15.

THIS book considers predation on the basis of evidence obtained in the wild, a welcome change from the generation of endless equations. The title notwithstanding, it is largely confined to freshwater zooplankton communities. Here, changes in the level and kind of predation can have dramatic effects on species composition and may even lead to complete restructuring. It is to these phenomena that Zaret addresses himself. Unfortunately the outcome is disappointing.

Two types of predators, gape-limited and size-dependent, are recognized. These, in effect, are vertebrates and invertebrates respectively. The logic of referring to fishes, much the most important vertebrate predators on the zooplankton, as gape-limited in this context is puzzling, for it is admitted that gape limitation is at most "of very brief duration in the fish's life and can be ignored for most considerations". To regard hemipterous insects, the only invertebrates so classified, as gape-limited seems equally odd.

The ways in which animals react to increased predation or pressure from different kinds of predators are described.

CHROMOSOMES TODAY

VOLUME 7
EDITED BY
M. D. BENNETT
M. BOBROW
& G. M. HEWITT

George Allen & Unwin announce the publication of the proceedings of the Seventh International Chromosome Conference held at Oxford in 1980. The volume provides up-to-date short reviews of a wide range of chromosome topics by leading authorities. It will be invaluable to all hard-pressed biologists who wish to keep abreast of recent research in the field. The main subject areas covered are: chromosome organisation at molecular level; chromosomes in meiosis; population cytogenetics; chromosomes and speciation; chromosomes and plant breeding; chromosomes and malignant change; induced chromosome variation and human cytogenetics.

0 04 575021 1 £16.00

George Allen & Unwin,
PO Box 18, Park Lane,
Hemel Hempstead, Herts HP2 4TE

NA 581

Circle No.29 on Reader Enquiry Card.

Bring Your Library Up-to-Date With This ANS Edition ...

INTERNATIONAL INSTRUMENTS FOR NUCLEAR TECHNOLOGY TRANSFER

...$49.00

Edited by L. Manning Muntzing
(6 x 9 hardbound, 639 pages)

American Nuclear Society
555 North Kensington Avenue
La Grange Park, IL 60525 USA

Please take my order for ______ books.
☐ Payment in full enclosed.
☐ Bill me and I accept postage and handling charges.

Name ______
Org. ______
Street ______
City ______
State ______ Zip ______

Circle No.05 on Reader Enquiry Card.

Small size or inconspicuousness achieved by various means enables some planktonic animals to evade fishes, and the possibility of doing so by vertical migration is also considered. Invertebrate predators can sometimes be thwarted by other means, as by protective spines or shapes that are difficult to handle, by escape reactions or simply by large size. Complexities arise when both types of predator are active: an anti-fish stratagem may increase vulnerability to a copepod. On the basis of case histories, somewhat subjective models are described which predict reactions to various kinds of predation and are shown to be in reasonable agreement with certain real life situations. Nature is, however, almost infinitely diverse and some categorical statements are incorrect. Thus *Bosmina* is not "dominant only when gape-limited predation is intense" — I know a score of fishless water-bodies where *Bosmina* dominates the open water, often alone — and it is nonsense to say that *Daphnia magna* (not truly planktonic) and *D. pulex* cannot co-exist. My records reveal *D. pulex* as the most frequent conspecific associate of *D. magna.*

And so to the factual errors, of which there are so many one hardly knows where to begin. The author does not know much about copepods or he would not refer to "large conspecifics" — adults do not moult — or ask "what function is served by their compound eyes?", organs they do not possess. Cyclopoids do not stab prey from behind with daggerlike mandibles. No mere quibble this: to understand protective devices we must know how the predator attacks. The statement that copepods "will attack neither very large nor very small prey" is also erroneous; cyclopoids that attack animals much larger than themselves, including small fishes, can be reared on protozoans. Considering its importance in the theories developed, size is indeed treated very carelessly. Some species of *Ceriodaphnia* attain more than twice the cited maximum size; the grotesque full-page figure of *Holopedium* bears the wrong scale; and to suit a model three daphnids are lumped together "because they are within a range of 1.0 mm", forgetting that in these small animals this reflects a difference of 125 per cent.

Other animals are equally abused. The full page figure that purports to be a *Chaoborus* larva (indexed as the phantom midget!) is only a head with part of the gut attached; the figure labelled *Bythotrephes longimanus* is *Polyphemus pediculus*; cladocerans not rotifers produce ephippia; stomatopods are not pistol shrimps nor notonectids water striders, and so one could go on. Carelessness, including a host of mis-spelled names, sometimes provides amusement — otter densities of 20 to 30 per m^2 must be as remarkable a sight as fishes "feeding on the mollusc's feet".

Revised and re-written this could be a useful little book: as it stands it cannot be recommended. □

Geoffrey Fryer is a Senior Principal Scientific officer at the Windermere Laboratory of the Freshwater Biological Association.

Physics and chemistry of membranes

Ora Kedem

Basic Principles of Membrane Transport. By S.G. Schultz. Pp.204 Hbk ISBN 0-521-22992-8; pbk ISBN 0-521-29762-1. (Cambridge University Press: 1980.) Hbk £14, \$22.50; pbk £4.95, \$8.50.

ALTHOUGH various treatises dealing with general aspects of membrane transport or membrane function are available, this book is a valuable addition to the literature. It is a unified presentation, starting from first principles and including discussions of some of the more important transport systems in the living organism. Nevertheless, it is appropriate that the title does not include any "bio-" key word because the author provides the physicochemical foundation for studying biological membranes, and does not attempt a summary of the detailed and rapidly growing physiological and biochemical information.

The structure of the book is logical and not historical; thus, for example, the GHK flux equation is thoroughly discussed in the chapter on isothermal diffusion, where it belongs. In the chapter on some principles of electrophysiology, the concept can then be applied in context. The brief summaries, at the end of some chapters, state any conclusions that one can draw from the relatively simple treatment and stress their limitations. Indeed, a feature of the book as a whole is that assumptions and arguments are clearly stated. While a formal quantitative account is given in enough detail to be readily followed by the reader, emphasis is on the physical significance of the results, and in many cases the physical meaning of steps in the derivation.

With this recommendation to both students and researchers in the field, a note of reservation is necessary. The desire for intuitively appealing explanation has led in some cases to inaccurate, or at least non-rigorous, statements. However, considering the logical development of the book as a whole — and the detailed references — these should not lead to any serious misunderstandings. □

Ora Kedem is Professor in the Department of Membrane Research at The Weizmann Institute of Science, Rehovoth, Israel.

Isotopes and geology

Gunter Faure

Handbook of Environmental Isotope Geochemistry. Vol.1, *The Terrestrial Environment*, A. Edited by P. Fritz and J.Ch. Fontes. Pp.532. ISBN 0-444-41780-X. (Elsevier Scientific: 1980.) Dfl.185, \$90.25.

THIS is the first of a proposed set of five volumes dealing with applications of isotopic studies to a wide variety of problems, ranging from the geochemistry of forest soil to igneous petrology. The first volume on the terrestrial environment gets off to a shaky start with a preface and an introduction by the series editors. They justify their ambitious undertaking with the claim that isotope geology reached its present importance in the earth sciences only in the context of geochemistry, hydrogeology and environmental geology. Applications of isotopic studies to dating of terrestrial (or extra-terrestrial?) materials and to palaeoclimatic studies are judged to be of limited practical importance; this point of view may explain why isotopic studies of interest "to economic and hardrock geologists" are relegated to the fifth volume. Yet, in spite of several factual errors, the introduction does contain much useful information on the systematics of isotope fractionation that may assist the reader in understanding technical details of the following twelve, more substantive, contributions.

The first five chapters are devoted to presentations of isotope fractionation in different terrestrial environments including the atmosphere, groundwater, ice and snow, and geothermal systems. These are followed by seven chapters dealing with more specific topics. All of the contributions were written by experts in their respective fields and contain a wealth of detailed information regarding isotopic and geochemical processes in different terrestrial environments. The articles by Deines on reduced carbon, Osmond on uranium disequilibrium in ground water, Savin on mineral–water interactions and Krouse on sulphur compounds in the environment are especially well organized and concisely written. Several others are a bit of a chore to read because they are long and discursive.

This book will be useful primarily to scientists in search of information about isotopic and geochemical studies when their own research impinges on the natural processes occurring on the Earth's surface. Each chapter is attended by a lengthy list of references that range over several decades from the early 1950s up to and including 1978, and there are detailed author and subject indexes which enhance the value of the book as a source of information. □

Gunter Faure is Professor of Geology at Ohio State University, Columbus, Ohio.

nature

4 June 1981

[illegible] Che[illegible]etry Library
[illegible] COLLEGE
[illegible] Upper [illegible]

15/2/82.

Madness, responsibility and the law

The law is necessarily an awkward compromise between the principles of justice and the needs of society. This simple truth has been illustrated all too vividly in the past few weeks by yet another sensational murder trial, in the genre that appears to recur every so often in England, and which on this occasion has led to the imprisonment of a Mr Peter Sutcliffe for a period of thirty years, after his admission that he had murdered no fewer than thirteen women over a period of several years. Mr Sutcliffe, when arrested, had confessed to his series of crimes but, at his trial, pleaded not guilty to the charge of murder, on the grounds of "diminished responsibility", but guilty of the lesser charge of manslaughter. The court declined to accept this version of events. The trial (before a jury) was an opportunity both for a rehearsal of the sombre details of Mr Sutcliffe's particular line in the technique of murder and for an examination of his state of mind at the time. Psychiatric specialists who had examined him told the court that Mr Sutcliffe suffered from "paranoid schizophrenia", and for their pains were required to spend several hours in the witness box at the Old Bailey speculating on the question whether Mr Sutcliffe could have learned his schizophrenic symptoms from his wife. In the end, the jury found Mr Sutcliffe guilty, the judge in the case called his crimes "cowardly" and Mr Sutcliffe was sent off to gaol.

The importance of the Sutcliffe case extends beyond what appears to have been its immediate function — providing British society with a cathartic ending to the long and incompetent hunt for a notorious murderer and satisfying the curiosity of the prurient. There are legal questions of peculiarly British interest about the validity of pleas of guilty to a lesser charge on the grounds of diminished responsibility. On this occasion, the British Attorney-General, as the country's principal law officer, had been willing to accept this plea so as to spare the relatives of the murdered women the pain of sordid publicity. (Now that capital punishment has been abolished, convictions for murder and manslaughter carry similar penalties, so that accepting Mr Sutcliffe's plea did not entail a promise of leniency.) In the event, the court declined to accept the compromise and the trial centred on the question of whether Mr Sutcliffe's responsibility for his actions had been undermined by his psychiatric condition. There are plainly enough intricacies in this decision to keep lawyers' dinner tables animated for years to come. The more important and durable issues have not so far been taken up.

The most immediate difficulty is not legal but social and medical. Mr Sutcliffe's case for supposing that he was not fully responsible for his awful actions was that he had the symptoms of paranoid schizophrenia, in particular the hallucination of hearing voices. The chief issue at the trial was whether Mr Sutcliffe might have simulated this and other possible symptoms of schizophrenia. The prosecution, the defence and even the court appear to have taken it as self-evident that if the diagnosis were correct, Mr Sutcliffe's bizarre actions would be accounted for and the plea of diminished responsibility made acceptable.

The fallacy in this line of argument is clear. People who suffer from schizophrenia are unhappily all too common, while paranoia is a common but not invariable symptom. Especially since the introduction of neuroleptic drugs as a palliative for schizophrenia, people diagnosed as paranoid schizophrenics have been living in the community, leading lives often no more aberrant than those considered as within the normal range for people free from psychiatric illness. Mr Sutcliffe had not sought treatment for his supposed illness (although, according to evidence during the trial, his wife had been treated for schizophrenia). Even so, there is no evidence to suggest that schizophrenia accompanied by paranoid behaviour predisposes towards sexual violence nor even that schizophrenia is more commonly represented than other types of psychotic illness in its causation.

All this goes some way to justify the court's view that the plea of diminished responsibility was unacceptable; evidence of paranoid schizophrenia, the argument seems to have been, is not by itself a proof of a diminished sense of responsibility. The hope now must be that all the talk there has been of the connection between schizophrenia and the likelihood of ending up in the dock at the Old Bailey will not too much unsettle the large numbers of people suffering from this illness who daily face crises less spectacular than Mr Sutcliffe's which are nevertheless potentially as catastrophic. The professional psychiatrists giving evidence last month, in their zeal to do the best for Mr Sutcliffe, may have overlooked the side-effects of what they were saying.

The deeper question is that of the relationship between psychiatric illness as now recognized in modern society and what used to be called madness. Where capital punishment has been done away with, there has usually been an intermediate phase in which the penalties of conviction for capital crimes have been moderated when it has seemed that the criminal could not reasonably be held responsible for his crimes, however dreadful. In Britain in the bad old days, juries in murder trials would frequently return verdicts of "guilty but insane" after careful instruction by the judge in the niceties of the McNaughton rules, intended to help decide whether a person carrying out a murder knew at the time the difference between "right and wrong". With the abolition of capital punishment, the need of this distinction has vanished, subsumed in two competing liberal principles — the view that people who commit what were once capital crimes are, by that test alone, out of their minds or mad, and the more subversive opinion (linked primarily with the name of R.D. Laing) that madness is simply a matter of the definition of normalcy, and thus a subjective matter. The Sutcliffe case has shown that British juries do not accept the first of these views and, fortunately, are probably ignorant of the second.

The result is an unhappy muddle, which cannot last much longer. The idea that the best way of treating serious psychiatric illness is to avoid, wherever possible, locking people up in hospitals has more than convenience and cost to commend it; institutionalization is, of itself, harmful even if on some occasions necessary. Helping people to lead more or less normal lives, however great the social cost and frequent the disappointment, does at least preserve the possibility of recovery from psychiatric illness. These are principles on which many societies are now agreed. Their practice is comparatively so recent, in the past quarter of a century, that a formal evaluation of the benefits would be premature. But it seems unlikely that the practice will be wholly abandoned of regarding the psychiatrically disabled as essentially no different from the rest of the population. One is, so to speak, presumed sane until proved otherwise.

Inconveniently, it follows that a diagnosis of psychiatric disability cannot logically be a synonym for diminished responsibility, whether in a sensational murder trial or in more ordinary circumstances. Psychiatric disability can of course be an extenuating circumstance, in a court of law or elsewhere, but that is a different matter. Both the prosecution and the defence at last month's unseemly trial at the Old Bailey would have been more

0028-0836/81/230367-02$01.00

persuasive if they had grasped that truth. The rest of us have still to wake up to the consequences of what is, in effect, the abolition of the concept of madness.

At first sight, the most obvious implications are in the legal system. People such as Mr Sutcliffe must be given the benefit of the doubt even if they are thus exposed to more serious penalties. The courts also need more helpful criteria for telling when the nature of an accused person's behaviour is such that full responsibility cannot be assumed. For the rest of us, the most telling consequence of the erosion of the concept of madness is not that R. D. Laing's subversive doctrines deserve respect, but that psychiatric illness deserves to be recognized for what it is — common and pervasive.

Nuclear politics

The past two weeks may turn out to have been more than usually significant for the development of nuclear power in the Western world. The fact that recent happenings are, in their way, self-contradictory, should not obscure their importance.

In France, the new president, M. François Mitterand, persuaded his first Cabinet last week that the plan to build another nuclear power station in Brittany should be put off at least until he has had a chance to organize the referendum on the future of nuclear energy in France promised during his election campaign: cynics are already saying that the gesture may win some votes in the coming elections for the Assembly without restricting his freedom to do what he chooses later on.

In Hamburg, almost the opposite has happened. After a year of furious opposition to the Brinkdorf nuclear power station, the Mayor of Hamburg has chosen to resign rather than to continue to threaten the already fragile stability of the government of Chancellor Helmut Schmidt, to whose political party the mayor belongs. In the United States, the Administration (which is no longer new) may not have the time, the energy or the inclination to fight Congress for the funds needed to finish off Clinch River, the first full-scale breeder reactor in the New World. In Britain, the argument (by the Monopolies Commission) has been more rudimentary: should a nationalized utility monopoly be entitled to charge its customers more than they would otherwise have to pay because it believes that it should be building nuclear power stations for "strategic" if not for economic reasons? With all these uncertainties in the air, nuclear engineers would be looking for quite different kinds of jobs if OPEC, at its meeting last week, had not agreed that the price of oil should be increased by a little less than the inflation rate (in dollars) over the past six months.

This curious conjunction of disparate events is a telling reminder that the time has come when the relationship between governments and nuclear energy should be changed. For the past thirty-five years (since the ending of the Second World War), it has been taken as read that nuclear energy is too important to be left to nuclear engineers. There are some grounds for this position. Obviously governments, small as well as big, have an interest in the proliferation of nuclear weapons — they either wish to stop it or to make sure they are not left out. Governments also have an interest not easily distinguishable from a duty to ensure that their electors will not be put at hazard by the operations of nuclear plants. It is also proper that governments should have a view on the place of nuclear energy in the evolving pattern of energy consumption. But none of this justifies the detailed concern of most governments for the intricacies of their nuclear power programmes.

If nuclear power had been an ordinary technology, government would long since have withdrawn from detailed involvement. Originally they were dragged in to regulate the use made of fissile material and to finance the high cost of reactor development. Although safety has been an issue from the start, it is only in the past decade that governments have been compelled by political pressure to consider the general question that the nuclear industry may be inherently too dangerous to be accommodated even in an advanced industrial society. President Mitterand's response is politically but not economically prudent — to halt the further development of nuclear power in France, and to hold a referendum; he knows that he will need the votes of as many French interests as possible to win a compliant Chamber of Deputies later in the year. The fact that it is the inverse of the 1979 response of Mrs Thatcher's then newly elected British government to general concern about nuclear energy suggests the unpalatable prospect that policies on nuclear energy are now politically partisan — right-wing governments build reactors, left-wing governments cry halt. That would be a recipe for doing nothing.

So there is an urgent need to rescue governments, as far as possible, from their political responsibility for nuclear power. How is this to be done? The prospects are not as hopeless as they sometimes seem. In the regulation of fissile material, the technical problems are well understood and the risks of failure keenly appreciated. Even so, on the principle that shedding administrative responsibility for nuclear matters is wise, governments should make fuller use the international safeguards that exist or could be created. The cost of nuclear development raises more serious difficulties. Reactors of the types now in commercial service are sufficiently well understood, technologically and otherwise, for decisions about their further exploitation to be left to electricity utilities and made on strictly commercial grounds. (This is not always simple, as shown by the recent criticism of the British Central Electricity Generating Board by the Monopolies Commission and by the Reagan administration's flat refusal last week to help with the cost of cleaning up at Three Mile Island.)

Utilities however are not in a position to support fast reactor development or experiments in thermonuclear fusion. It is proper that taxpayer's money should be used instead. It is essential that governments should distance themselves from the use eventually made of successful developments, insisting that success does not predicate widespread use. (Successive United States administrations have done this consistently, but the Clinch River reactor has also been hindered by doubts about cost and design).

But none of this will absolve governments of responsibility on the general question of safety. If a substantial part of an electorate is persuaded, even if mistakenly, that the civil nuclear enterprise is inherently unsafe, a democratic government has no choice but to grasp the issue firmly — if it wants to be re-elected. The question to be answered now in France is whether repeated declarations that nuclear energy is an abomination are to be taken at face value.

That there should be local opposition to the building of nuclear plants and that arguments about safety should arise is not surprising. The same happens when chemical manufacturers look for new sites. Would-be operators of nuclear plants must be prepared to allay reasonable doubts on safety. Then it is also right that planning authorities (whose authority derives in the last resort from central government) should arbitrate and if necessary override local opinion — otherwise, national interests must be perpetually subservient to local interests. Technically this is what will be happening next year, when the Central Electricity Generating Board will apply to build a pressurized water reactor plant in Suffolk. The snag is that the ultimate arbiters in the planning process will be members of the British government — the Ministers of the Environment and of Energy. Should they not also distance themselves from the process by saying in advance that they will accept whatever recommendations the judicial enquiry makes?

The wider question of whether there should be nuclear energy at all is much harder to resolve. President Mitterand's referendum is likely to create as much confusion as those in Sweden and Austria in the past two years. The overriding difficulty is that it is never possible to be certain that those declaring themselves against nuclear energy have been fully seized of the counter-arguments — the need for alternative sources of energy whatever they are, for example. On issues such as nuclear energy, decision by referendum is almost always bad decision — governments must shoulder the responsibility. What else are governments for?

NIH censure for Dr Martin Cline

Tighter rules for future research plans

Washington

Dr Martin J. Cline, the California scientist who has admitted using recombinant DNA molecules in unapproved experiments with human patients, has been rapped sharply over the knuckles by the US National Institutes of Health (NIH). The director of NIH, Dr Donald Fredrickson, announced last week that he has accepted the recommendation of an investigatory committee that Dr Cline be required to obtain special permission from NIH in future to carry out research using recombinant DNA techniques, the first time such action has been taken against an individual investigator.

Dr Fredrickson also agreed that review committees assessing future grant applications from Dr Cline be provided with a detailed account of his admitted transgressions in attempting to treat patients with thalassaemia. However, the investigatory committee, which described his actions as a clear violation of NIH guidelines covering recombinant DNA research, has not suggested stronger action.

Rather it has proposed — and Dr Fredrickson has accepted — that the directors and advisory councils of the three institutes from which Dr Cline at present receives grants totalling about $600,000 a year should decide whether any of his current support should be withdrawn. Dr Cline's sentence therefore rests essentially in the hands of his scientific peers, a precedent-setting situation which is being closely watched by groups such as the President's Commission for the Study of Ethical Problems in Biomedical Research.

Dr Cline resigned in February from his position as chief of the division of haemotology–oncology at the University of California, Los Angeles (UCLA), although he remains on the university faculty as professor of medical oncology. His resignation followed the admission that, in conducting experiments with patients suffering from β-thalassaemia in Italy and Israel last summer, he had transplanted bone marrow cells whose genetic material had been altered by recombinant DNA techniques. This was despite the fact that he had earlier told the hospital authorities that his proposed therapy would not involve such types of cell.

Dr Cline told the NIH committee that he was unaware that strict federal regulations requiring prior approval of experiments involving human subjects was required for any experiment carried out by university researchers, whether in the United States or abroad. The UCLA committee subsequently told NIH that the definitions of collaborating institutions contained in documents distributed to research staff "are not as clear as they might be".

It was the unauthorized use of recombinant DNA molecules, however, that made the charges against Dr Cline more serious. In written evidence to NIH investigators, he says he took samples of cloned genes intending to perform *in vitro* studies of β-thalassaemic marrow in Italy and, if possible, organize clinical trials in Israel and Italy.

On the last day of his visit to Israel, he received permission from authorities at the Mount Scopus Hospital of the Hassadah medical complex to use purified genes in experimental therapy with a 21-year-old girl suffering from β-thalassaemic major, a blood disease rare in the United States but more common in the Mediterranean region. Having originally intended — and received the permission — to carry out the experiments with unaltered genetic material, Dr Cline decided to subject bone marrow removed from the posterior iliac crest to a DNA-mediated gene transfer technique using recombinant genes which had been planned for the *in vitro* studies.

A similar procedure, in which the treated bone marrow was replaced in the patient in an attempt to stimulate the production of healthy red blood cells, was performed four days later on a 16-year-old girl with the same disorder at the University Polyclinic in Naples, Italy. Describing the Israeli experiment, Dr Cline later told NIH that he had decided to use the recombinant genes on medical grounds "because I believed that they would increase the possibility of introducing β-globin genes that would be functionally effective, and would impose no additional risk to the patient, since it was known that pieces of DNA are efficiently linked in all combinations once they are taken into cells".

The medical authorities at the Israeli hospital, however, who had taken great pains to check that Dr Cline's proposed experiment did not involve the use of genetically altered cells, were upset at discovering what had taken place. Dr Cline's collaborator at Hadassah, Dr E.A. Rachmilewitz, denied any knowledge that recombinant DNA was being used and investigations carried out at UCLA confirmed that Dr Cline had acted on his own.

Accepting Dr Cline's letter of resignation, the dean of the UCLA School of Medicine, Dr Sherman Mellinkoff, accepted that no known harm had been done to the two patients, and that the experiments might prove helpful in treating others. But he added "It is also true that the freedom to conduct experiments of benefit to mankind is jeopardized by failure to act in accord with the relevant regulations in these circumstances".

Both Dr Cline and the university have declined to comment on the report of the NIH committee and Dr Fredrickson's subsequent decision. However, in a letter to NIH written in January, Dr Cline says "I greatly regret my decision to proceed with the use of recombinant DNA molecules without first obtaining permission from the appropriate committees", adding that "I exercised poor judgement in failing to halt the study and seek appropriate approval".

David Dickson

Social scientists take over ANZAAS forum

Brisbane

The 51st Congress of the Australian and New Zealand Association for the Advancement of Science (ANZAAS), held at the University of Queensland in Brisbane (11–15 May) was publicized as a "stocktake for science", but it turned out that the association itself was most under scrutiny. The New Zealand contingent has threatened to withdraw and go independent. The prominence of social science has increased confusion about the role of ANZAAS, which will be added to at the next congress by the introduction of new sections on law and robotics.

For basic and applied scientists, ANZAAS has become a forum for retrospective reviews and personal contacts but no longer an occasion for the announcement of original research results. The specialist professional societies of Australia have now completely taken over that job. Thus more than 600 members of the Australian Society for Microbiology attended its annual conference in Canberra in the previous week but only a few dozen people turned up to the microbiology section at ANZAAS.

ANZAAS's own 10-times-a-year journal *Search* publishes only a minute handful of the 800-plus papers presented at each congress. There is little professional discipline over the speakers — papers are not vetted before delivery and too many speak from notes. Although many speakers took their responsibilities seriously, the overall standard was just too patchy. Presentation of a paper at ANZAAS may provoke valuable discussion, but it does not guarantee publication or significant enhancement of scientists' reputations.

For social scientists and some humanities, however, ANZAAS has become a welcome place for the communication of new research results and ideas. Given the science-and-society approach of some of the more traditionally scientific sections of ANZAAS the social sciences have now become dominant at this, Australia's largest annual academic meeting.

The organizers of the Brisbane congress

0028-0836/81/230369-02$01.00

latched on to the popularity of energy as a topic for endless talk and took "Energy and equity" as the theme. Australia is going through an excitable phase over energy matters. Australia's huge coal reserves are proving immensely attractive to foreign interests. Oil shale deposits, notably in Queensland, were the subject of wildly optimistic predictions of national self-sufficiency in oil, until the dominant Exxon partners pulled back on grounds of massively increased cost estimates.

Unfortunately, the symposia which offered well-known people a chance to air their well-known views did not significantly advance the professional, popular or political understanding of the facts and issues. Full publication of these addresses is unlikely, and the brief mentions they received in newspapers caused little impact. The organizers failed to capitalize on the large media contingent present with the energy theme buried in a snowstorm of unrelated items. The paper which attracted most news reporting and comment — even a newspaper leader — was a sociological study of the sub-culture of nudist colonies.

ANZAAS congresses suffer from a lack of continuity. Each meeting is organized in a different city by a totally different group. ANZAAS also seems to have no collective memory, so that the failings of one meeting are repeated. Its central organization is weak because it depends too much on volunteers — its membership subscription income remains static and it receives no government subsidy.

The enrolment of 2,300 at the Brisbane congress was disappointing, being right on the financial break-even point. ANZAAS will therefore now have to struggle financially for another year in the hope that next year's meeting in Sydney will attract enough delegates to lift its spirits and wipe out its deficit.

For all its faults, however, ANZAAS remains a potent force in Australian (not New Zealand) academic life, which could realise its potential if only it could get its act together. This year's congress came only two weeks after Prime Minister Malcolm Fraser announced a series of cost-cutting measures recommended by a group of senior ministers, popularly known as the "Razor Gang". The Razor Gang slashed into the supposed "fat" of publicly funded science and did so to an extent probably greater than any other single area. An undiscriminating cut of 2–3 per cent in the staffing levels of all public service operations will hit CSIRO hard, though this has as yet not been publicly recognized. Another act of the Razor Gang — the imposition of fees for second degrees in universities — is the cause of nationwide protests by staff and students alike.

Fears for the future health of Australian science were all the talk at ANZAAS, but only privately. Not a murmur was heard publicly about these problems. ANZAAS 1981 may well be remembered as an opportunity lost. **Peter Pockley**

Mitterand's new ministers

On your Marx

With the all-important French legislative elections looming on 14 and 21 June — which will determine whether the new socialist president, François Mitterand, has real power or not — it may be unwise to read too much into the events of the past few days. But Mitterand's much-vaunted commitment to science seems in fact to be one to technology — albeit technology with a socialist mask.

Mitterand has appointed Jean-Pierre Chevènement, a 42-year-old left-wing intellectual and founder of the Centre for Socialist Studies, Research and Education (the broadly Marxist CERES), as Minister of State for Research and Technology. The other principal contender for the job, the scientist and director of the Institut Pasteur, Professor François Gros, has been appointed as scientific advisor to the Prime Minister, M. Pierre Mauroy.

Chevènement is a great admirer of the highly-centralized Japanese ministry for industry and technology, MITI; and as a Minister of State he will be a member of Mitterand's powerful inner cabinet of five. He will rank higher even than the Minister of Industry, M. Pierre Joxe (who, incidentally, also calls himself a Marxist).

Chevènement and Joxe are now locked in battle over who will control which of France's many scientific and technological institutions. The Delegation Génerale pour la Recherche Scientifique et Technique (DGRST), which draws up guidelines and coordinates research throughout the government, but has only a small autonomous budget will be Chevènement's of right; and it seems he has also won control of the principal body funding basic science in France, the Centre National de la Recherche Scientifique (CNRS), from the ministry of education. The Centre National pour les Etudes Spatiales (CNES) which controls France's scientific and technological work in space, may also be a major prize from the ministry of industry, and another could be the Agence Nationale de Valorisation de la Recherche (ANVARD), which promotes innovation in French industry. The control of the medical research body (INSERM) and the agricultural research council (INRA) is also in question.

Even the atomic energy authority (the CEA) might fall to Chevènement, but once the dust of Mitterand's energy policy has settled, the science and technology minister will be seen to be strongly pro-nuclear (see box). He is also greatly interested in nuclear weapons (he backs a French strategic nuclear capability), and in military research.

Where is socialism — let alone Marxism — in all this? It comes partly in nationalization: Mitterand has promised to nationalize nine big industrial companies, to give greater state control of major technical sectors of the French economy (this will be M. Joxe's preoccupation). It comes in promised efforts towards less centralized, more democratic control of technological decision making (Mitterand speaks of a referendum on nuclear power, for example). And it may come in Chevènement's professed concern for the impact of new technologies on employment and working conditions. His attitude to the problem of short-term contracts for young researchers and technicians, which has flared up regularly in the universities, CNRS, INSERM and INRA in recent years is not yet clear. **Robert Walgate**

French face energy questions

Energy policy in France is entering a rather confused period. The new government has cancelled the bitterly contested plans for a nuclear power station at Plogoff in Brittany but has announced the start-up of two new 900 MW stations at Gravelines and Tricastin. And when the outgoing government in its dying days gave the go-ahead for doubling the size of the spent fuel reprocessing plant at Cap de la Hague there was no demur from the already-elected Mitterand.

So where lies the new French government's energy policy? There is no energy minister in the new government. The President, it seems, would like to keep this difficult card in French politics to himself. And so far, in fact, Mitterand's pre-election promises and his first actions as President have been consistent. He offered only a "pause" to reconsider nuclear power, in which reactors under construction (like Tricastin and Gravelines) would continue to completion; Plogoff has not begun. And Cap de la Hague has been so inefficient it needs refurbishment to cope with future spent fuel.

But there is disagreement in his cabinet about what to do next. Some say Plogoff is cancelled; others, that it is merely postponed. Mitterand may call some kind of referendum (and if he does, is likely to find the French in favour of nuclear power) or he may opt for a parliamentary inquiry, through his promised committee on technology assessment, or he may hold a public inquiry.

But barring accidents like Three Mile Island, France's nuclear programme will probably survive the storms, because the country needs the energy, and soon. Despite the vociferous opposition at Plogoff last year, the electorate of the region in fact voted overwhelmingly for pro-nuclear candidates in the first round of the recent presidential elections. The (Giscardian) member of parliament for the region complained last week "where will it be possible at sure cost, to find 5,200 MW for Brittany by 1990? And another 4,200 jobs?". **Robert Walgate**

0028-0836/81/230370-01$01.00

Corporate largesse

More tax breaks?

A tax incentive measure designed to stimulate corporate support for scientific research in US universities has taken a first step towards congressional approval. A comprehensive tax bill sponsored by Representative James M. Shannon has passed a House Ways and Means Committee panel on its way to the full committee.

Representatives from major universities, including the presidents of the Massachusetts Institute of Technology and Stanford, have testified at House hearings along with prominent industry members in support of the legislation. They hope that the Shannon bill will offset the expected losses from the Reagan budget by attracting more than $500 million to university research next year.

Three sections of the tax incentive plan would specifically benefit universities:

- A 25 per cent tax credit for contributions from corporations to a "research reserve fund" that would guarantee against shortfalls in research funding. This money, which could be up to 5 per cent of a corporation's income, can remain in the fund for four years drawing interest, while the taxpayer receives more credit for additional money added.
- A 25 per cent tax credit for money contributed to any form of research and development at universities or private laboratories if the total contribution exceeds the average amount given by that corporation over the previous three years.
- Tax deductions to corporations giving charitable contributions of scientific or educational equipment, based on the market value rather than production cost.

The "research reserve fund" would actually give businesses a larger tax credit because a contributing corporation would receive an additional deduction of up to 46 per cent when its money was transferred from the fund to a specific university.

Opposition to the bill comes from those who believe it would entail a major loss of federal revenue (around $207 million), but the White House has indicated in informal discussion with House members that it would definitely consider supporting it.

Besides bringing more corporate dollars to American academic research, which at present makes up less than 5 per cent of the total given to universities, supporters of the bill say it would foster closer ties between university scientists and applied research.

Passage of the bill may, however, be hindered by the lack of a unified academic voice. Some smaller universities feel that it will hurt them because most money will be directed towards the larger institutions. But insiders say the bill has a bright future because Senate Finance Committee chairman, Senator Robert Dole, has expressed great interest in its passage.

Michael D. Stein

Yugoslavian university unrest

Cuts feared

The University of Pristina, the third largest university in Yugoslavia, may be cut to about one fifth its present size following the political unrest of the last three months, which began in the university.

Pristina is the capital of the autonomous province of Kosovo which forms part of the Republic of Serbia but enjoys considerable local self-government. It ranks as one of the less developed areas of Yugoslavia, and hence, over the last few years, has benefited considerably from federal development schemes. Under such a scheme, the University of Pristina has grown during the last decade to a student body of almost 50,000, instead of the 10,000 originally planned, an expansion deemed necessary by Belgrade for "the overall development of Kosovo and the advancement of national equality". (The population of Kosovo is predominantly Albanian).

However, the growth in numbers was not matched by a commensurate expansion of student facilities, and on 11 March there was a small demonstration of students demanding improvements in hostel accommodation and canteen food. The local and university authorities seem to have over-reacted when restoring order and almost inevitably, the demonstrations escalated and became more political in nature. By the end of April, demonstrators were demanding outright secession of Kosovo from the federation and even a link-up with Albania; cultural and scientific exchange programmes between Yugoslavia and Albania were subsequently cut.

Although by now the demonstrations had spread far beyond the student body, the university was still regarded as their focus. Accordingly, on 18 May, Pazajit Nusi, vice president of Kosovo Assembly's executive council, and Dr Imer Jaka, Secretary for Education, Science and Culture, resigned. The same evening there was another on-campus demonstration, and the next day the executive council ordered the university and all other institutions of higher education in Kosovo to close for the vacation.

The university party organization, in the meantime, had set up its own action committee. According to one report more than 500 members of the philosophical faculty have been expelled from the Party, and four persons suspected of organizing the demonstrations have been expelled from the university. Also, various abuses in the university structure have been uncovered. At the same time,the action committee was highly critical of the Belgrade authorities and media for exaggerating the situation, and was particularly critical of Milan Daljevic, Executive Secretary of the (Federal) Central Committee Praesiduim, who had spoken of the University of Pristina as "a fortress of nationalism" and had alleged that not a single professor had made any political statement condemning "enemy actions". An official protest has been sent to the Central Committee Praesidium.

To offset the incomplete, and often incorrect, media coverage, the action committee called for a complete dossier on the unrest to be prepared.

With so few jobs available for existing graduates and 55,000 young people now receiving higher education in Kosovo, the suggestions from Belgrade that the university be cut back to its original 10,000 places and funding switched to other educational needs, poses a real threat to the University of Pristina. So far the action committee has made no explicit reference to the suggestion, but has merely suggested that future admissions be tailored more closely to "social needs".

Vera Rich

US security worries

No compromise

Washington

In a move reflecting deep-seated disagreement over how far national security arguments should influence the procedures of basic research, scientists on two advisory committees to the National Science Foundation (NSF) have criticized proposals by the American Council on Education (ACE) that papers containing potentially sensitive results in computer research be voluntarily submitted to national security authorities for review before publication.

The proposal was made by a study group on public cryptography set up by ACE at the suggestion of the National Security Agency (NSA). This followed a series of incidents in which NSA had sought to restrain the dissemination of unclassified data on the grounds that they could threaten national security by assisting the development of techniques to make and break codes.

NSA already accepts in principle the proposals of the study group, and has outlined to several professional organizations and their scientific journals the group's conclusions under which suggestions for changes, deletions or publication delays would be made on manuscripts submitted to the agency. ACE announced last month that "the intelligence and academic communities have agreed on guidelines for voluntary review before publication of computer cryptography articles". However, members of the mathematical and computer sciences advisory committee of NSF have said that voluntary submission to pre-publication review represents a "direct and serious threat to NSF's charter of furthering basic scientific research". A report, prepared by a subcommittee under the chairmanship of Dr John Guttag, says that the conclusions of the ACE study

0028-0836/81/230371-02$01.00

group "reflect a view of 'national interest' and 'national security' that is derived largely from concerns related to the government's need for military and diplomatic security". Other factors which it says should be considered include the competitive needs of the US information processing industry, and the need to protect individuals from unwarranted invasions of privacy.

The two advisory committees decided last autumn to investigate the role of NSF in supporting cryptological research, following an incident in which a computer research scientist who, having applied to NSF for a grant, was told that NSA was interested in supporting his research — even though he had not approached the security agency. In general, the committees endorse the policy outlined at the time by NSF deputy director Dr Donald Langenberg.

However, the committee's report takes issue with Dr Langenberg's remark that imposing more stringent reporting requirements on scientists should be viewed "simply as a change in administrative practice". Rather, it says, any change in current practice should be considered a significant shift in policy which could represent "serious interference with the conduct of research".

The committees also object to the letter being sent to those who are awarded cryptography grants notifying them that they should seek prior review by NSF of any results which might have national security implications. The report believes this would adversely affect research and proposes instead that a description of the research results should be forwarded to NSF or to any other government agency designated by the foundation — at the same time as its dissemination to other scientists or publication.

The NSF report points out that NSA has proposed spending between two and three million dollars a year on unclassified research — "We doubt that this much money could be spent productively supporting academic work on cryptography and cryptoanalysis, and imagine that NSA will end up funding and taking an active interest in a relatively wide variety of research projects". And it warns of the potential dangers of NSA becoming the principal source of funding for cryptology research as the Defense Advanced Research Projects Agency has done in fields such as artificial intelligence and systems research.

The Association of Computing Machinery (ACM) is also concerned over moves in Congress to give the federal government greater powers to control the publication of scientific or technical data. In particular, ACM has criticized a bill introduced by Representative Charles E. Bennett of Florida which would extend existing controls on "militarily critical" technologies to cover not only hardware, but any related technical information.

Dr Peter Denning, the president of ACM, feels that the bill would extend to many areas of research the type of "born classified" restrictions now applied to nuclear research under the terms of the Atomic Energy Act, and used as the basis for the (unsuccessful) prosecution last year of the *Progressive* magazine for publishing details of how a hydrogen bomb works.

The Reagan administration's attitude to this whole area is uncertain. Mr Lawrence Brady, nominated to head the Commerce Department's office responsible for export controls, promised a Senate committee last month that a new East–West trade policy, expected from the White House in a few weeks, would include rigorous new criteria for screening critical technology.

Mr Brady's appointment, however, has worried representatives of several multinational companies who feel that his zeal in restricting high technology exports to the Soviet Union may prove costly to them. The Administration is trying to define an acceptable compromise between the high priority it has given "national security" arguments, and the freedom being demanded by both technology exporters and the academic community. The sharp division between the ACE study and the NSF advisory committee indicates that such a compromise may still be far away.

David Dickson

French to make Arabsat

The French aerospace firm la Société Nationale Industrielle Aérospatiale (SNIAS) has won the coveted contract for "Arabsat" — a clutch of three communications satellites with an option on a fourth, for the Middle East — in competition with British Aerospace and the Hughes Corporation. But hidden behind SNIAS is Ford Aerospace, which will get 55 per cent of the $134.3 million deal. Until recently Ford was under an Arab boycott, and the company clearly thought it politic to work with a French partner, French–Arab friendship having been so strong under President Giscard d'Estaing.

British Aerospace are understandably downhearted. The firm has so far sold satellite systems only to the European Space Agency (ESA), and it saw Arabsat as a potential breakthrough into new markets.

But British Aerospace still have orders for seven satellites worth £100 million, which will keep them busy until 1985. The company is building five communications satellites for ESA, each capable of carrying 12,500 telephone conversations and two television channels. These are bigger than the Arabsats which will be designed for 8,000 telephone lines. British Aerospace are also building two maritime "Marecs" satellites for ESA, the first due to be launched by Ariane in October.

Meanwhile SNIAS is crowing just as British Aerospace would have done if it had won Arabsat. It is the first time a French aerospace firm has won the place of prime contractor for a foreign satellite system. SNIAS now have their eye on Australia, Columbia, Mexico and Brazil.

Robert Walgate

UK agricultural research

Change at the top

The recent appointment of two chief scientists to the British Ministry of Agriculture, Fisheries and Food is expected to increase the ministry's influence on research supported by the Agricultural Research Council. But there is to be no formal change in the relationship; this has been ruled out by the Secretary of State for Agriculture, together with the possibility that more money should be transferred from the council to the ministry for spending on commissioned research.

These developments help to clarify more than a year's uncertainty about the future relationship between the two organizations. After the House of Commons Public Accounts Committee recommended in July 1979 that the ministry should have greater control over the council's research, a committee under the chairmanship of Sir Brian Hayes, permanent secretary at the ministry, was set up to decide what should be done.

One of its conclusions was that the Rothschild principle, whereby almost half of the council's research is commissioned by the ministry, is to remain intact. Another was that the ministry could exert more influence if it coordinated its own science policy more effectively. To that end, the former chief scientist, Professor B. G. Weitz, has been replaced by two chief scientists, Professor Frank Raymond for agriculture and horticulture and Dr George Elton for fisheries and food.

The idea is that two chief scientists will enable the ministry to make better use of the Agricultural Development and Advisory Service (ADAS), which accounts for almost half of the ministry's staff and advises farmers and officials on technical aspects of agricultural and horticultural practice. Professor Raymond will report to ADAS and formulate an appropriate policy for research. Thus the influence of ADAS within the ministry will also be enhanced. Dr Elton will concentrate on fisheries and food research. Both chief scientists will sit on the Agricultural Research Council.

Precisely how the new arrangements will affect the council remains to be seen. But the outcome contrasts sharply with the situation for medical research: last October, the Medical Research Council won back most of the money transferred to the Department of Health, so doing away with the Rothschild principle.

Judy Redfearn

0028-0836/81/230372-01$01.00

Origin of Species

Premises, premises

Origin of Species is the latest in the series of new permanent exhibitions scheduled by the British Museum (Natural History). Under any circumstances, it would be an important exhibition, lying at the heart of what the museum is trying to explain. Unless there is some central thread of theory linking together louse and lousewort, *Archaeopteryx* and albatross, bushbaby and bushman, the exhibits that surround us in the museum are of no greater significance than a collection of postage stamps — interesting for their variety, occasionally beautiful, but intellectually sterile. But circumstances today make us look even more closely at such an exhibition, for evolutionary theory seems to be assailed ever more fiercely by its enemies, and to be only half-heartedly defended by its supporters.

Let's start, though, with the presentation and packaging, which is what will (or will not) attract the visitor in the first place. The Origin of Species exhibition is mounted in the western wing of the premises designed by Waterhouse. Spacious, elegant and well-windowed, it is a beautiful foil for the exhibition, which is arranged in eleven sections, each presenting a different aspect of the theory of evolution by natural selection: the species, competition, variation, inheritance and so on. With the inevitable few minor reservations, I thought that this system worked very well: each concept is comprehensible and is clearly and attractively explained. Both diagrammatic and real-life presentations are used, and are cleverly intermingled. There is also a natural selection computer game, designed to appeal to the Space Invaders generation. Finally, there are four mini-theatres in which film-loops explain different aspects of the theory.

Understanding the details of Darwin's theory is, however, only one part of what we scientists hope our museums will do for us. In the face of the organized pressures of religious and mystical sects, evolutionists need some organization to represent *their* views, no less fervently held, as cogently as possible. Not that it should descend to the half-truths and double-talk of political propaganda. But it should suit the terms of its message to those who will listen to it, rather than blunting its edge with the hair-splitting logic-chopping of the philosophy of science. Here, I fear, the BM (NH) has failed us.

"The Survival of the Fittest is an empty phrase; it is a play on words. For this reason, many critics feel that not only is the idea of evolution unscientific, but the idea of natural selection also. There's no point in asking whether or not we should believe in the idea of natural selection, because it is the inevitable logical consequence of a set of premises." "The idea of evolution by natural selection is a matter of logic, not science, and it follows that the concept of evolution by natural selection is not, strictly speaking, scientific." "If we accept that evolution *has* taken place, though obviously we must keep an open mind on it . . . ." "We can't prove that the idea is true, only that it has not yet been proved false." "It may one day be replaced by a better theory, but until then . . . ." These are all quotations from the film loop in which the present status of the theory of evolution is explained to a layman by a scientist. If this is the voice of our friends and supporters, then Creation protect us from our enemies.

Of course, to a scientist, evolution and natural selection are "only" theories — but how much real doubt is implied by that statement? The philosopher of science may utter his customary caution "Anything you say will be taken down; it will only be falsifiable, and can never be proved true". But these theories have withstood over a century of testing, by a great variety of methods unimaginable to Darwin. These methods not only all support a theory of evolution in general fashion, they combine to support a single, precise version of the relationships of living things. For example, palaeontology, biogeography, serology, protein sequencing etcetera, *all* suggest that man is most closely related to the African apes, rather than each type of evidence pointing to a different branch of the tree of life. The significance of Mendel's work was not merely that it did *not* support Lamarck (as the film-loop tells us), but also that it was exactly what Darwin's theory demanded and had, to that extent, effectively predicted (which the film-loop does *not* tell us). How many biologists really think that, one day, they will open *Nature* and find a new, completely different theory of evolution?

Equally, if we say that Darwin's theory may one day be replaced by a "better theory", do we really mean anything *completely* different? In fact, of course, few biologists consider it remotely possible that the great variety of evidence that supports this theory could suddenly be re-arranged to provide an even more persuasive support for a creationist theory. But that is not all. We don't even think that it could support a dramatically different *scientific* theory, in the way that earlier observations of the heavens were transformed from being compatible with an Earth-centred Universe to demonstrating a Sun-centred Solar System. Genetic drift or neo-Lamarckism may lead to changes of emphasis or provide explanations of the exceptional or the rare, but they show no signs whatever of taking over the central ground of our science. Though no biologist can deny the possibility that God created man, few would doubt that, if He did so, the mechanism that Darwin discerned was the one that He chose to use. It is the strength of conviction of working biologists that needs to be accurately translated to the man in the museum gallery, not the technical caveats with which we remind ourselves of the potential weaknesses of more recent or less wide-ranging theories. And let us not forget that Dr Roger Miles, head of the museum's Public Services Department, is quoted as saying that their exhibitions are aimed at the average 15-year-old, not at scholars.

Lest I be accused of losing my sense of proportion, let me repeat that I found most of the exhibition appealing, satisfying and clear. If I take issue at length with one small part of it, it is because I consider that part to be crucial. This film-loop will be studied primarily by those few people who are really interested in trying to find out precisely how the world of science views this most important and wide-ranging explanation of the phenomena of the living world, and they may judge this against the less dispassionate appeals of such groups as the creationists. Just as war may be too serious to be left to the military, so the presentation of science is too crucial to the scientist to be left wholly at the mercy of the populist presenter. **Barry Cox**

Huxley on *Nature*

At the celebrations of the centenary of the move of the British Museum (Natural History) to the building it now occupies in South Kensington (London), the president of the Royal Society, Sir Andrew Huxley, devoted the major part of his speech to an attack on *Nature* for giving publicity to criticism of the museum (of which Sir Andrew is now a trustee) in recent months.

Sir Andrew complained that *Nature* has "printed more than 30 letters centering around attacks on the scientific thinking of members of staff of the museum" and said that the correspondence, especially that relating to cladistics, would have been "unintelligible to 99 per cent of the readership of *Nature*".

The "obscurities and irrelevances" of the debate were "in cloud-cuckoo land". Sir Andrew said that *Nature* published such letters for the same reason that *The Times* of London devoted "so much space to the Sutcliffe trial" (a recent sensational murder trial in Britain) — "their readers enjoy these things".

Sir Andrew also accuses *Nature* of having published a leading article about the museum "on the basis of a few words" taken out of context, and of not having published "a brief reply from the museum". Sir Andrew warned *Nature* that "crying wolf will drive its readers into disregarding its future warnings" and those "wounded by these articles" should ignore them.

0028-0836/81/230373-01$01.00

CORRESPONDENCE

Apes or angels?

Sir — The description of the punctuated equilibrium model as Marxist[1], although correct, obscures its real importance in the thinking of east coast radicals: the theory is at root anti-racist. As the debate between Rose[2] and Dawkins[3] made clear, the radical scientific opposition to racism requires a denial that there is any genetic variation of any significance from place to place within the human species. As there is a considerable amount of geographical variation in anthropometric characters and in the frequencies of blood groups and other biochemical traits, anyone who wishes to espouse both the neo-Darwinian synthesis and this form of scientific anti-racism must adopt the argument either that the genes affecting behaviour just happen not to be subject to any effective geographical variation in their distribution[4], which is uncomfortably close to special pleading, or that behaviour is not under genetic control. As this second proposition is known to be untrue for other animals, there is an urgent need for a general theory that will rescue scientific anti-racism from sophistry.

The punctuated equilibrium theory[5,6] holds that all species make qualitative "leaps" at the time they originate. This can be applied to humans as meaning that the particular and special "leap" made by our species was largely in freeing the functioning of the brain from genetic influences: no matter how gene-distributions may vary geographically, or hormones between the sexes, the psyche is entirely buffered against these and is influenced only by the environment[7]. Further, human races are nowhere near to speciation and hence, not having "leapt", are fundamentally "the same". The punctuated equilibrium theory is thus the modern radical version of Disraeli's famous question and answer "Is man an ape or an angel?". Gould is on the side of the (evolutionary) angels.

However, a belief in the absolute moral value of one's own theological (or scientific) position carries its own danger of personal damnation; this is the lesson of all witch-hunts. Although Gould[8] would like us to believe that he is the potential victim this, in view of recent history, is a piece of *chutzpah*. Within the academic community it is those scientists whose theories could be conveniently labelled "racist" or "right-wing" who have been subjected to unofficial but nonetheless unpleasant persecution (the assault on Eysenck).

Most scientists do not, as Gould[8] implies, resent the implication that their theories may be founded in their socio-political world-view: they simply regard this obvious fact as irrelevant. The argument that a theory is incorrect merely because the proponent has ulterior motives for holding it has been known as a species of false logic since the ancient Greeks. However, as radical scientists just do not accept the fact–value separation which Dawkins[3] and other liberals use to defend their position, but rather adopt the Marxist view that moral statements are founded in objective reality (the only true science is that which serves the good of the people)[9], they interpret the moral statement that racism is wrong as implying the scientific statement that sociobiology is wrong. And as the Soviet ambassadors in Islamic countries have been trying to explain recently, it is important to distinguish a progressive force which swims with the moral tide of history from a reactionary imperialist force which swims against it. Perhaps the same distinction applies to witch-hunts.

J.R.G. Turner

Department of Genetics,
University of Leeds, UK

1. Halstead, L.B. *Nature* **288**, 208 (1980).
2. Rose, S. *Nature* **289**, 335 (1981); **290**, 431–432 (1981).
3. Dawkins, R. *Nature* **289**, 528 (1981).
4. Cavalli-Sforza, L.L. & Bodmer, W.F. *The Genetics of Human Populations* (W.H. Freeman, San Francisco, 1971).
5. Stanley, S.M. *Macroevolution, Pattern and Process* (W.H. Freeman, San Francisco, 1979).
6. Lewin, R. *Science* **210**, 883–887 (1980).
7. Futuyma, D.J. *Evolutionary Biology* (Sunderland, Massachusetts, 1979).
8. Gould, S.J. *Nature* **289**, 742 (1981).
9. Brecht, B. *Leben des Galilei* (Frankfurt, 1955).

It wasn't me

Sir — You recently published a letter on genetic determinism from someone in Massachusetts with a name remarkably similar to mine. My acquaintances all seem to think that I wrote that letter, so I wish to state publicly that this is not the case. His first name is spelled "Isadore".

Isidore Nabi

University of Chicago,
Illinois, USA

Unsöld on Einstein

Sir — I feel that the article "German physics in row about Einstein" (*Nature* 16 April, p.535), dealing with a lecture and a publication of mine, needs some correction. It states that I began my "attack on Einstein" at a symposium held last May. In fact, this refers to a lecture, "Evolution of cosmic, biological and metal structures", which I gave on 25 April 1980 at a physics colloquium at the University of Marburg. With regard to the evolution of mankind, I mentioned the alarming problem of weapons and actions of mass destruction (gas warfare, the atomic bomb, napalm, atomic and neutron rockets, and, of course, Hitler's genocide), for which highly intelligent people were responsible. In the discussion which followed, no one, from an audience of some 200, said anything about Einstein. However, weeks later, a Marburg professor wrote to many physicists, complaining bitterly about my lecture.

As to my article in *Physikalische Blätter* dealing primarily with the Einstein celebrations of 1979, I, of course, fully acknowledge the great discoveries of Einstein, but I strongly disagree with the many statements which attempt to show him as a great politician or even a kind of saint. And I disagree still more with a prominent speaker who claimed that, compared with Einstein, Planck's political attitude was "childlike".

Being no historian, I took almost all my biographical information from the book *Einstein, the Life and Times* by Ronald W. Clark, a prominent British historian of science. And the president of the Society for History of Sciences, Professor F. Krafft, stated in a letter of 29 April 1981 that, as far as he could see, my article contained no false statements or quotations.

The publication of my article was solely the responsibility of the five editors, one of whom was Professor Rollnik. Any attempt to shift the editorial responsibility to others would be quite unfair.

As to the reaction of others, I would like to emphasize that I received many letters of support from prominent scientists, while it seems Professor Rollnik received most of those which disagreed with me. Several of them seem to use the term "Nazi" in exactly the same way the Nazis used the term "White Jews", in order to outlaw all theoretical physicists in Germany!

Albrecht Unsöld

Institute of Theoretical Physics,
University of Kiel, FRG

Sir — Under the West German press laws I am the editor responsible for the *Physikalische Blätter*, the journal of the German Physical Society, and would like, in this capacity, to say a few clarifying words on the article "German physics in row about Einstein" (*Nature* **290**, 535; 1981).

I intentionally did not involve Professor Rollnik, the president of the German Physical Society and co-publisher of the journal, in the discussions on the publication of Professor Albrecht Unsöld's manuscript as I do not consider it a good thing for the president of a scientific society to decide on the views, political or otherwise, of the society's members. A society's journal, which reflects only the opinions of the president, would indeed be a questionable enterprise.

I therefore much regret that Herr Rollnik has been rebuked. Why did the critics not write to Professor Unsöld or to the editorial staff?

Unsöld has received many letters agreeing with him and this endorsement largely accords with my reasons for deciding to publish the article: in no way did Unsöld want to enrich the Einstein biographies. He simply wanted to take Einstein (and Haber) as examples to illustrate that the trend towards increasingly more effective means of mass destruction, however well-intentioned they are subjectively, ultimately only enlarges the catastrophe. Unsöld destroyed the *idol* "Einstein" in order to get at the main theme of his article, the increasingly dangerous disproportion between man's moral capacities and his intellectual faculties. I am probably not the only one to share this concern with him.

Unsöld speaks expressly of the great "tragedy" of nationalism and how this also involved great natural scientists. This was not meant to discredit, but nor was it intended as antisemitism. Surely it was also tragic that nuclear weapons were developed by leading scientists of a people Hitler had horribly plagued. Nevertheless, this is of little help to us in Central Europe if these weapons are to be used in our countries. Not even the 30,000 Jews who still, or once again, live in Germany could hope to be spared.

Karl Kromphardt

Physikalische Blätter,
Bad Honnef, FRG

0028-0836/81/230374-01$01.00

NEWS AND VIEWS

Geophysical controversy in the Gulf of Aden

from R.B. Whitmarsh

THE concept of seafloor spreading has been used successfully to unravel the history of the major ocean basins in terms of varying speeds and directions of spreading of lithospheric plates including, in some cases, periods of slow or negligible spreading. There remains, however, the problem of understanding the processes whereby a continent is split apart to produce a new ocean. For this reason there has been interest for some years in studying nascent or very young oceans, of which the best examples are probably the Gulf of Aden and the Red Sea. These represent the westward extension of the Carlsberg Ridge from the north-west Indian Ocean, and pass northwards into the sinistral Gulf of Aqaba/Dead Sea shear or transform fault.

The oceanic nature of the Gulf of Aden crust was established in 1966 by Laughton[1] on the basis of seismic refraction data. The presence of a mid-ocean ridge running down the centre of the Gulf was convincingly demonstrated by the presence of a central rough zone with large magnetic anomalies, offset by NE–SW fault zones, and the existence of a chain of earthquake epicentres and high heat-flow observations. Laughton did not attempt to date the oceanic crust which, on the basis of east–west faults exposed ashore, appeared to be younger than late Eocene to early Miocene. However, the existence of a central topographical rough zone, flanked by crust covered by 0.5–1.5 km of sediment, caused him to suggest two phases of opening of the Gulf.

Normally geophysicists date the oceanic crust by subjectively matching a sequence of seafloor spreading magnetic anomalies with a section of the geomagnetic reversal time scale which complies with the constraints offered by independent geophysical and geological information, such as palaeontological dates based on core samples. The difficulties and limitations of this method, when the anomaly sequence is short, have recently been amply demonstrated by the publication of two different spreading histories for the Gulf of Aden[2,3].

Cochran[3], after Laughton *et al.*[4] but with more data, favours a single phase of NE–SW seafloor spreading for the past 10 Myr. In both these papers the data base consisted of a set of variously oriented profiles from all parts of the Gulf. On all profiles, but particularly in the western and central Gulf, the anomaly amplitude decreases markedly away from the central anomaly. Fairly convincing correlations of calculated (synthetic) and observed anomalies can be obtained out to anomaly 5 (10 Myr BP) east of 45°E. West of this longitude the anomalies only fit well out to 3 or 4 Myr BP. Cochran's synthetic profiles do not convincingly reproduce the amplitude variations which occur away from the spreading axis.

Girdler and colleagues[2,5] favour a more complex history of the Gulf of Aden involving at least two (and possibly three[2]) phases of seafloor spreading. Their data set consists of 16 magnetic profiles, obtained in a single survey and carefully oriented along the spreading direction, which cross the spreading axis between 43° and 46°E in the western Gulf of Aden. Like Cochran they believe that spreading has been continuous for about the past 5 Myr and they too could not identify anomalies older than anomaly 3 generated by the spreading axis west of 44°20′E. The difference in their interpretation lies in the proposal that there was an earlier phase beginning 23.5 Myr ago, followed by an 11.5 Myr hiatus in spreading from 16 to 4.5 Myr BP. Also, using geophysically plausible arguments about the decrease of magnetization with age in the upper crust, they were able to model successfully the amplitude variations in their survey area. Girdler and co-workers have attempted to choose a spreading history which is consistent with other geological and tectonic information; in particular, their two-phase model fits the last two major phases of transform motion along the Dead Sea rift as identified by Freund.

The disagreement between Cochran and Girdler's group lies simply in the interpretation of a 50 km-wide strip of magnetic anomalies, each side of the central spreading zone, which is at least 5 Myr old. As Cochran shows, profiles calculated from the two models are qualitatively very similar out to his oldest identified anomaly. This surely indicates the futility of trying to resolve the problem by modelling seafloor spreading anomalies. Outside the above strip Girdler's group were able to model other anomalies for a further 50 km but these were of low amplitude and were evident on only some of their profiles.

As the oceanic lithosphere spreads away from a mid-ocean ridge, it cools and contracts and its upper surface (the acoustic basement on reflection profiles) follows a characteristic age versus depth curve[6]. Therefore it is implicit in the Girdler 11.5 Myr hiatus that a step should appear in the acoustic basement of about 450 m. Unfortunately, the roughness of the acoustic basement in the Gulf of Aden is such that an offset of this size in the mean level of the basement is barely resolvable and difficult to demonstrate convincingly. On the other hand, Cochran quotes depths of the basement producing his oldest anomaly which are apparently consistent with a continuous spreading history for the past 10 Myr.

As well as regional geological information concerning the ages of tectonic events on adjacent plate boundaries, there is the evidence of a Glomar Challenger borehole at site 231, some 150 km from the spreading axis. Here in the southern part of the central Gulf of Aden fossiliferous chalk about 13 Myr old was found embedded within the basaltic basement[7]. This age, while consistent with Cochran's interpretation, conflicts with an estimated basement age of 25 Myr BP from

R.B. Whitmarsh is at the Institute of Oceanographic Studies, Wormley, UK.

---

1. Laughton, A.S. *Phil. Trans R. Soc.* **A259**, 150 (1966).
2. Girdler, R.W., Brown, C., Noy, D.J.M. & Styles, P. *Phil. Trans. R. Soc.* **A298**, 1 (1980).
3. Cochran, J. *J. geophys. Res.* **86**, 263 (1981).
4. Laughton, A.S., Whitmarsh, R.B. & Jones, M.T. *Phil. Trans. R. Soc.* **A267**, 227 (1970).
5. Girdler, R.W. & Styles, P. *Nature* **271**, 615 (1978).
6. Sclater, J.G., Anderson, R.N. & Bell, M.L. *J. geophys. Res.* **76**, 7888 (1971).
7. Fisher, R.L. & Bumce, E.T. *Init. Rep. DSDP* **24**, 17 (1974).
8. Roeser, H. *Geol. Jahrb.* **13**, 131 (1975).
9. Searle, R.C. & Ross, D.A. *Geophys. J. R. astr. Soc.* **43**, 555 (1975).
10. Vine, F.J. *Science* **154**, 1405 (1966).
11. Noy, D.J. in *Tectonics and Geophysics of Continental Rifts* (eds Ramberg, I.R. & Neumann, E.R.) 279 (Reidel, Holland, 1978).

0028-0836/81/230375-02$01.00

the Girdler model. It is conceivable, however, that the basalt represents an intrusive sill, thereby post-dating the underlying crust.

Because horizontal, and particularly strike-slip, motion across the East African Rift is believed to have been negligible in the past few tens of millions of years, a corollary of any model of opening for the Gulf of Aden is that essentially the same history must fit observations in the Red Sea. In the latter, the magnetic field also displays high-amplitude axial anomalies which can be correlated out to about 3–5 Myr BP[8-10]. Here too, Girdler's group favours a two-phase spreading history[11] but this has yet to be matched precisely with their latest history for the Gulf of Aden[2].

The published evidence does not favour one model more than the other. The difficulties of interpreting short magnetic profiles are evident, although a systematic magnetic survey recently conducted in the eastern Gulf of Aden where the spreading rate, and therefore the resolution of magnetic profiles, is greatest could yet provide conclusive evidence. The problem provides a salutary lesson on the necessity of combining the best geophysical and geological data and the difficulties which can arise if evidence from a single observational technique is used to study the Earth. □

The transmission and control of gonorrhea

from Robert M. May

In the United States, roughly one million cases of gonorrhea are reported each year, and the Center for Disease Control suggests that between two and three million new cases may actually arise annually. An estimated 10–17 per cent of women with gonorrhea develop pelvic inflammatory disease (PID), which is the most common complication of the disease and is the leading cause of sterility in women. It is claimed[1] that direct and indirect costs related to PID and associated ectopic pregnancies exceeded 1.25 billion dollars in the United States in 1979.

For diseases such as measles, chicken pox or influenza, people recover relatively rapidly and thereafter become immune. Thus the average fraction of the population that is infected (the prevalence) cannot increase indefinitely, but rather saturates at a level where most contacts made by an infected person are with individuals who have acquired immunity. For these diseases, the prevalence therefore depends on the rate at which new susceptibles appear in the population (by birth or immigration), and it tends to be necessary to have large aggregations of people — typically 500,000 or more — in order to maintain the infection.

In common with most other sexually transmitted diseases, gonorrhea has characteristics causing its epidemiology to be quite different from the above infections. First, only sexually active individuals who could be infected by their contacts need to be considered as candidates in the transmission process. In contrast with simple models for infections such as measles or the common cold, a doubling of the population density does not double the overall rate at which new infections are produced by infected individuals (unless one makes the bizarre assumption that doubling the population density doubles an individual's typical rate of acquisition of sexual partners). Second, many infected people, especially women, are essentially asymptomatic and therefore do not seek treatment. Since individuals do not recover spontaneously from gonorrhea until a relatively long time has elapsed, such infected persons will remain infectious until they do receive antibiotic treatment. Third, little or no immunity appears to be acquired on recovery from gonorrhea, so that most treated individuals rejoin the susceptible class as soon as the effects of antibiotics wear off. All this adds up to an infection that is admirably well adapted to persisting in small, low-density aggregations of humans.

The first mathematical model to cater specifically to these features of the epidemiology of gonorrhea was developed by Cooke and Yorke[2]. Subsequent studies by Reynolds and Chan[3], Hethcote[4], Constable[5] and others have attempted to estimate the parameters in models involving sets of coupled differential equations. Much of this work is reviewed by Yorke, Hethcote and Nold[6] (hereafter YHN), who offer many important insights; their presentation is aimed at public health workers and is deliberately uncluttered with mathematical details.

YHN emphasize that, for gonorrhea, saturation effects arise when infectious individuals contact individuals who are already infected from some other source. They call this 'preemptive' saturation, and suggest that "preemption appears to be the only saturation factor that can affect a disease such as gonorrhea for which infection does not confer immunity".

Turning to the data for gonorrhea in the United States, YHN estimate the annual incidence to be 2.6 million cases per year, and the average duration of infection for women and men to be 55 days. This corresponds to a prevalence of around 400,000 cases at any given moment. Further assuming the 'at-risk' population of sexually active people to be around 20 million, YHN arrive at the result that approximately 2 per cent of the at-risk population is infected at any one time; changing the estimate of the at-risk population to 10 or to 50 million does not affect the essentials of the subsequent argument.

Given that only around 2 per cent or so of the contacts of an average infective person are themselves already infected, preemption cannot be a significant factor if the at-risk population is considered as one homogeneous, freely mixing group. Thus either gonorrhea must currently be far from equilibrium and its prevalence rapidly growing in the United States (which does not appear to be the case), or else we must, in YHN's words, "separate the population at risk for gonorrhea into many subgroups, by dividing according to sex, age, race, sexual practices, number of sex partners and other relevant factors". These authors argue, for example, that a 'core' of around half-a-million people having high prevalence (of the order of 20 per cent) would have substantial preemptive effects, and could account for the observed data. Essentially what is happening is that each infective individual in the core population, on average, infects more than one susceptible individual during the course of his infection, while each infective individual outside the core, on average, infects less than one susceptible; at equilibrium, these two rates combine to produce an overall average of one new infectee for each current infection.

The existence of such a core of very sexually active transmitters has immediate implications for the design of control programmes. If the core individuals could all be identified and kept free of gonorrhea (for example, by use of some hypothetical vaccine), the disease would die out of the population, because its basic reproductive rate in the remaining non-core population is less than unity.

In Hethcote, Yorke and Nold's most recent work[7], they subdivide the population according to sex, the level of sexual activity among the at-risk population (core and non-core) and whether the infections are symptomatic or asymptomatic. A set of eight differential equations is used to model the dynamic interplay between these groups and the consequences of various kinds of control strategies are explored. In these models, the probability of transmission during a sexual intercourse from an infectious female to a susceptible male is estimated[8] to be 0.2–0.3, while the corresponding

---

Robert M. May is Class of 1877 Professor of Zoology at Princeton University.

0028-0836/81/230376-02$01.00

probability of transmission from an infectious male to a susceptible female is around 0.5–0.7; since a new sexual encounter typically involves more than one sexual intercourse, the net female-to-male transmission probability may be around 0.5, and the male-to-female around 0.9, per encounter. These detailed values are of less significance in the models than is their ratio, namely around 1:2 (female-to-male versus male-to-female).

Broadly, Hethcote, Yorke and Nold[7] consider three control strategies. Strategy AW involves the general screening of females, by testing cultures taken from women who use certain designated health facilities. The corresponding strategy AM does the same thing for males. Strategies BW and BM trace females and males, respectively, to whom an infection has been spread by a diagnosed infected individual. Strategy CW identifies infectious females by the men they have infected; strategy CM correspondingly identifies infectious males by their women infectees. In short, strategy A takes a slice in time, B pursues the chain downstream, and C pursues it upstream.

Although the quantitative details depend on the specific parameters chosen in the various models, strategy C is always more effective than A, which in turn is markedly superior to B. YHN explain this by observing that asymptomatics are more important transmitters than symptomatics because they are infectious for a longer time, and that core persons are more important transmitters because they contact more people while they are infectious. Strategy C has a high likelihood both of identifying very active core individuals and of identifying asymptomatics; strategy A, although relatively ineffective in locating core people, tends to find asymptomatics; while strategy B is not particularly effective in finding either core individuals or asymptomatics. Within any one strategy, finding infectious males by M controls is generally at least as effective at reducing prevalence levels as is finding infectious females by W controls.

Hethcote, Yorke and Nold[7] sensibly conclude by stressing that "the relative merits of the various control procedures depend not only on the effectiveness per discovery in reducing prevalence, but also on the cost of discovering an infective by the procedures". If field studies can be designed to evaluate the cost in dollars, or at least the relative costs, for the discovery and cure of one person by each of the general strategies, the way will be opened to the rational design of 'cost effective' programmes. "Different cities or states might come to different conclusions based on the differing nature of their populations[7]." □

---

1. *United States Morbidity and Mortality Weekly Report* **28**, 605 (1980).
2. Cooke, K. L. & Yorke, J. A. *Math. Biosci.* **16**, 75 (1973).
3. Reynolds, G. H. & Chan, Y. K. *Bull. Inst. int. Statist.* **106**, 264 (1975).
4. Hethcote, W. H. *Math. Biosci.* **28**, 335 (1976).
5. Constable, G. M. *Bull. Inst. int. Statist.* **106**, 256 (1975).
6. Yorke, J. A., Hethcote, H. W. & Nold, A. *Sex. Trans. Dis.* **5**, 51 (1978).
7. Hethcote, H. W., Yorke, H. A. & Nold, A. *Math. Biosci.* (in the press).
8. Wiesner, P. J. & Thompson, S. E. *Disease-a-Month* **27** (5), 1 (Feb., 1980).

---

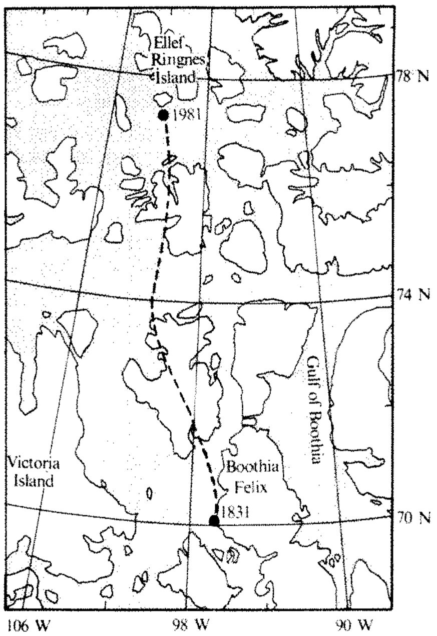
150 years of the North Magnetic Pole

from David R. Barraclough and S. R. C. Malin

On June 1st, 1831, Commander James Clark Ross "fixed the British flag on the spot, and took possession of the North Magnetic Pole and its adjoining territory, in the name of Great Britain and King William the Fourth"[1]. The spot referred to was at 70° 5′17″ N, 96° 46′45″ W where the mean value of magnetic dip, observed over nearly 24 hours, was 89° 59′. This was, as nearly as could be determined with the instruments then available, the position of the North Magnetic Pole (where the dip would be exactly 90°). Commander Ross thus became the first man knowingly to set foot on (there was an eskimo settlement nearby) the North Magnetic Pole.

James Clark Ross was not a stranger to Arctic exploration. After entering the Royal Navy at the age of 12 he served under his uncle, John Ross, for six years. In 1818 he was a member of the expedition, led by John Ross, which began the series of nineteenth century assaults on the North-West Passage. For the next nine years the younger Ross was almost continuously in the Arctic, serving under Parry on his four famous expeditions.

In 1829 John Ross organized an expedition, financed by Felix Booth, the gin distiller, to make yet another attempt to find the North-West Passage. As his second-in-command he chose his by now very experienced nephew. It is not clear how importantly the discovery of the position of the North Magnetic Pole figured in the original plans for the expedition, but it was ultimately to prove the major achievement, since the attempt to find the North-West Passage failed. In fact, having spent the first winter frozen in on the east coast of Boothia Felix (now the Boothia Peninsula), the expedition ship *Victory* was able to move only about 8 km northwards during the next two years in an unsuccessful attempt to escape from what had proved to be a dead end. During this enforced stay in Boothia Felix, the younger Ross led several sledging expeditions across and along the peninsula. He also made a series of geomagnetic observations from which he estimated the probable position of the North Magnetic Pole to be within range of a sledging party. Accordingly, during the last summer in Boothia, James Clark Ross set out to discover the Pole's position and succeeded in doing so.



Path of the North Magnetic Pole, 1831—1981.

The story of how the expedition finally extricated itself and was almost miraculously rescued after a fourth winter in the Canadian Arctic cannot be told here. John Ross was knighted in recognition of his achievements and his nephew became the acknowledged leader in the field of polar exploration. James Clark Ross was himself awarded a knighthood on the return of his successful Antarctic expedition during which he came close to, but was unable to reach, the South Magnetic Pole.

The North Magnetic Pole has moved about 800 km north during the 150 years since its discovery, as indicated in the figure. Its present position[2] is about 77.3°N, 101.8°W.

Sir James Clark Ross's achievement, together with those of other British geomagneticians, is being marked in an exhibition 'The Earth is a Magnet' to be held in the Royal Scottish Museum, Edinburgh, from June 8 until the end of August.

David R. Barraclough and S. R. C. Malin are at the Institute of Geological Sciences, Edinburgh.

---

1. Report of Commander, now Captain, James Clark Ross, RN, FRS, FLS in *Narrative of a Second Voyage in Search of a North-West Passage, and of a Residence in the Arctic Regions during the years 1829, 1830, 1831, 1832, 1833* (Webster, London, 1835).
2. *Magnetic Declination Chart 1980.0*, Canada (Division of Geomagnetism, Department of Energy, Mines and Resources, Ottawa).

0028-0836/81/230377-02$01.00

The treatment of inborn errors of the urea cycle

from Isabel Smith

THE UREA CYCLE is the means by which mammals remove excess ammonia from the body. Every enzyme in the cycle (see the figure) is subject to inherited abnormality, thus giving rise to a group of rare diseases known as the urea cycle disorders[1]. With the exception of ornithine carbamyl phosphate transferase (OCT) deficiency, which is a sex-linked condition, the disorders appear to be inherited recessively. They are characterized by a tendency to accumulate ammonia, which may be severe enough to cause death within a few days of birth, or so mild that it is present only after a protein meal and remains symptomless throughout life.

Typical symptoms of ammonia intoxication include recurrent attacks of vomiting, lethargy and drowsiness, often precipitated by catabolic events such as infection, or by increased protein intake. Severely affected patients commonly progress to coma and death, often in the neonatal period. Mental retardation, epilepsy and other neurological abnormalities are frequent in the survivors. Levin and co-workers[2] reported beneficial effects of a reduced protein intake on heterozygous, female carriers of OCT deficiency, and since then a low-protein diet supplemented with extra fat and carbohydrate has formed a basis for treatment. Unfortunately these measures alone fail to prevent ammonia intoxication in patients with more severe types of illness, and in growing children prolonged use of a low-protein diet may cause deficiency of essential amino acids and thereby exacerbate nutritional problems arising from the feeding difficulties caused by hyperammonaemia.

Recent advances in therapy have come through a better understanding of the special nutritional needs of patients with urea cycle disorders, and through exploitation of improved methods of reducing intake, and increasing utilization and excretion, of excess nitrogen[3-7]. Life-threatening episodes of hyperammonaemia can be treated with peritoneal or haemodialysis before irreversible brain damage occurs[3,8,9]. In addition, it has been realized that not all symptoms are due to ammonia intoxication. Patients with arginase deficiency, for example, become severely retarded and have limb spasticity although they exhibit only slight hyperammonaemia. Their neurological abnormalities appear to be caused by arginine accumulation. A patient whose blood arginine levels were reduced by means of a low-arginine diet, which was started in the neonatal period, made normal intellectual and neurological development, in striking contrast to the typical symptoms present in an older untreated sibling[10].

Arginine has also become important in the treatment of other urea cycle disorders. It is the only amino acid of the cycle to be incorporated into protein. Normally it is formed in the cycle and is non-essential, but when a block occurs, an exogenous supply of arginine is needed for protein synthesis. In disorders other than arginase deficiency plasma arginine levels are low,

---

Isabel Smith is at the Institute of Child Health, University of London.

---

100 years ago

A SINGULAR CASE OF SHIPWRECK

THE wreck of the Danish mail steamer *Phoenix* which took place off the west coast of Iceland on January 29, was attended by rather unusual circumstances deserving of note. The vessel (about 450 tons burden) sailed with cargo and the mails from Copenhagen for Leith, the Faröe Islands, and Iceland, about the middle of January.

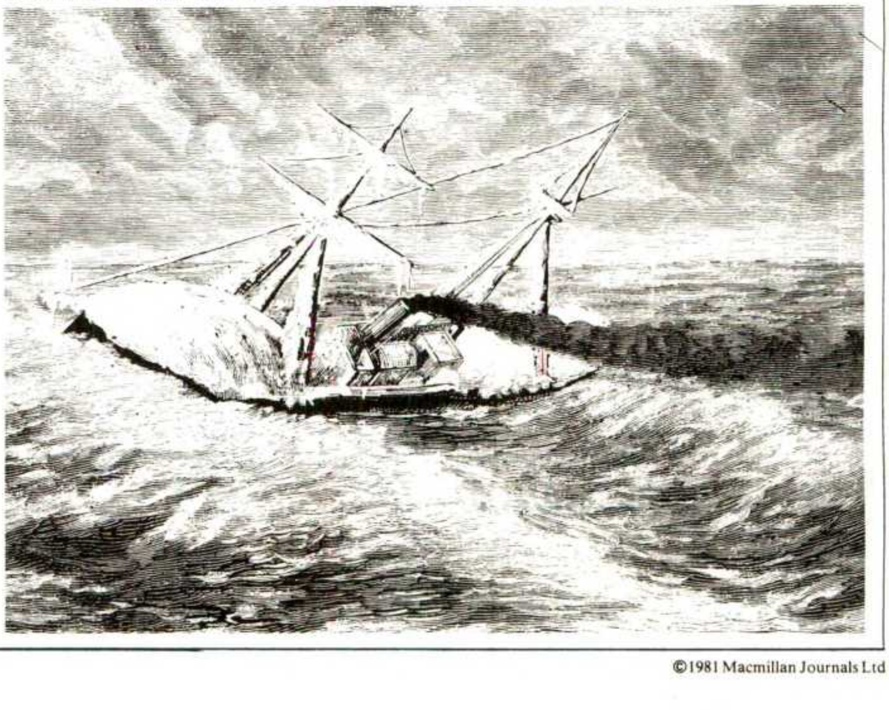
Nothing particular occurred until after leaving the Faröes, when she ran into a severe gale, which rapidly increased to a perfect hurricane, while at the same time the temperature fell to about 50°F. of frost (—18°F.). Such cold is not extraordinary in these latitudes in winter but fortunately it is seldom associated with very high winds. Under the circumstances in which the *Phoenix* was placed every sea that she shipped froze, and the deck soon became covered with a foot or two of solid ice.

As time passed on the continued action of the sea raised a perfect iceberg on the forward part of the vessel, while the showers of spray carried along by the steadily increasing gale covered the masts, yards, and rigging with an ever-thickening coating of ice. Two or three days passed without the least abatement of the storm, and then the half-smothered steamer went over on her beam ends. The crew succeeded in cutting away the masts, and she once more righted.

It however was clear, the gale showing no signs of breaking, that the relief was only temporary. The ice continued to form on the vessel, particularly about the forecastle, where, piled high above the bulwarks, and overhanging the sides, it threatened, by altering her trim, to raise the propeller out of the water.

Under these circumstances, on the morning of January 29, Capt. Kihl decided to run the steamer ashore while daylight lasted. At some distance from the land she struck on a sunken rock, and the crew, taking to the boats, only succeeded with the greatest difficulty in reaching the shore, saving nothing but their lives, the English mail, and a bundle of blankets which (when carried ashore) was found to be useless — frozen into a solid lump. Hardly one escaped more or less injury from the effects of the extreme cold to which they had been so long exposed. Capt. Kihl and the bulk of his crew soon after succeeded in getting to Reykjavik, and on April 13 they sailed in the sister steamer, the *Arcturus,* for Copenhagen. The officers and men of the wrecked vessel are of opinion that had Capt. Kihl not decided on the 29th to run the *Phoenix* ashore in daylight not a soul would have been saved, as the gale did not moderate for several days after; and the steamer, buried as it was under an enormous mass of ice, must have foundered in the night.

from *Nature* **24**, 2 June, 106, 1881

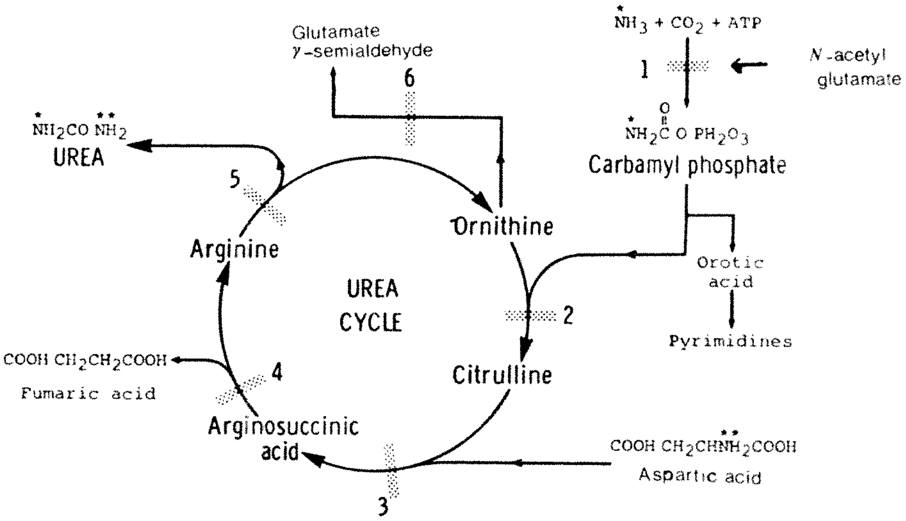


0028-0836/81/230378-02$01.00

contributing to growth impairment. A low-protein diet given without arginine supplements increases the tendency to arginine deficiency. Poor growth from any cause reduces utilization of nitrogen, thereby tending to increase ammonia accumulation. In arginosuccinic acid lyase (ASL)[6] deficiencies, lack of arginine causes not only growth impairment but also, by reducing the supply of ornithine (as discussed in detail below), contributes significantly to persistence of raised ammonia levels.

The role of arginine in ammonia homeostasis arises because it is the main precursor of ornithine which is required for uptake of carbamyl phosphate[11]. The potentially reversible ornithine amino transferase (OAT) step seems normally to proceed towards glutamate semialdehyde. In experimental animals arginine deficiency causes decreased intracellular ornithine levels and hyperammonaemia, even when the urea cycle itself is intact[11]. A similar situation arises in human patients with lysinuric protein intolerance, which is due to an inherited defect in the transport of lysine out of the cells of small-bowel mucosa, liver and renal tubules[12]. Lysine accumulates intracellularly leading to competitive inhibition of arginine and ornithine uptake and to intracellular ornithine deficiency. Administration of arginine intravenously rapidly abolishes the hyperammonaemia[13], and oral supplements of arginine, or ornithine, or, best of all, citrulline (because uptake of this amino acid is not inhibited by lysine), maintain normal ammonia levels over long periods when given in combination with a low-protein diet[12].

Although not, strictly speaking, urea cycle disorders, two different inborn errors of ornithine metabolism, both of which lead to high plasma levels of ornithine, illustrate relevant points about the role of arginine in ammonia homeostasis. One, OAT deficiency[14], is accompanied by progressive gyrate atrophy of the choroid and retina starting during the first or second decade of life. Patients are of normal intelligence. Plasma levels of ammonia, and its precursors glutamine and glutamic acid, are abnormally low probably because intracellular levels of ornithine, like plasma levels, are high. Treatment with a low-arginine diet restores normal plasma levels of ornithine, ammonia, glutamine and glutamic acid, and there is some evidence that progression of visual abnormalities is halted. The other disorder, which seems to be due to a defect in the transport of ornithine into mitochondria[15], is accompanied by severe mental retardation and marked hyperammonaemia, probably due to intracellular deficiency of ornithine in spite of the elevated plasma levels. As might be expected of such a disorder, patients derive no benefit from a diet which lowers plasma ornithine levels.



1, carbamyl phosphate synthetase (CPS). 2, ornithine carbamyl phosphate transferase (OCT). 3, arginosuccinic acid synthetase (ASS). 4, arginosuccinic acid lyase (ASL). 5, arginase. 6, ornithine aminotransferase (OAT). *One nitrogen atom enters the cycle at step 1; **one nitrogen atom enters the cycle at step 3. Reactions 1, 2 and 6 take place within the mitochondria, 3, 4 and 5 in the cytosol.

Lowered intracellular ornithine levels are likely to play a major part in the production of hyperammonaemia in arginase deficiency and in ASS and ASL deficiencies. From a study of an infant with ASS deficiency (also called citrullinaemia), Walser *et al.*[4] estimated that a minimum arginine supplement of 2.2 mmol per kg per day was needed to provide sufficient ornithine for carbamyl phosphate uptake and to allow for normal growth. Arginine supplements combined with a natural protein intake of 1.7 g per kg per day (just at the lower limit of normal for infants) maintained normal ammonia levels for several months; development proceeded normally. The infant died suddenly of hyperammonaemia, however, during an infection at the age of seven months.

An even more impressive effect of arginine supplements (2–5 mmol per kg per day) has been described in patients with ASL deficiency[6]. In this disorder both nitrogen atoms can enter the cycle normally, and be excreted as arginosuccinic acid, providing sufficient arginine is available (see the figure). Ammonia levels can be maintained in the normal range when 2 g protein per kg per day are given with arginine supplements. With smaller supplements of arginine[16], normal growth was maintained but control of ammonia levels was inadequate and the patient showed slow mental development.

Bachmann and Colombo have noted the excretion of large amounts of orotic acid (a derivative of carbamyl phosphate which is produced in increased amounts when the urea cycle is blocked beyond CPS, see figure) in the urine of patients with arginase deficiency, in spite of the mild degree of ammonia accumulation. In ASS and ASL deficiencies, in which ammonia levels are higher, much less orotic acid is excreted. These findings suggest that production of carbamyl phosphate may be increased in arginase deficiency and Bachmann and Colombo have proposed that this could be due to stimulation, by high arginine levels, of *N*-acetylglutamate synthetase, an enzyme responsible for production of *N*-acetylglutamate, which is the essential co-factor for CPS. Activity of CPS is directly proportional to the concentration of co-factor[18]. There is some *in vitro* evidence that arginine does increase *N*-acetyl synthetase activity[17]. If the hypothesis is correct, arginine deficiency could reduce ammonia disposal by reducing CPS activity, as well as by the mechanism discussed above, and arginine supplements might thus improve control of ammonia levels by stimulating CPS in OCT, ASS and ASL deficiencies and in partial defects of CPS.

Whatever the mechanisms involved, arginine supplements added to a reduced protein intake benefit many patients with urea cycle defects. In severe forms of CPS, OCT and ASS deficiencies, however, additional measures are needed. One approach is to replace natural protein by a mixture of essential L-amino acids[3]. This allows nitrogen intake to be reduced, stimulates nitrogen utilization for synthesis of non-essential amino acids, and ensures that intake of essential amino acid requirements for growth are met. Another similar approach, using a mixture of keto-acid analogues of essential amino acids[19], reduces intake, and stimulates utilization, of nitrogen even further. As threonine and lysine cannot be synthesized *in vivo* by transamination, and as the keto-acid derivatives of certain other amino acids are in short supply, a mixture of keto-acid analogues (leucine, isoleucine, valine, phenylalanine and methionine) is given with a mixture of essential L-amino acids (threonine, lysine, tryptophan, histidine), with arginine, to replace natural protein[20,21]. Patients with severe CPS, OCT and ASS deficiences, which have been treated in one of these ways, have remained healthy, with good control of ammonia levels, for many months.

A very different approach to the control of ammonia levels has been introduced by Brusilow *et al.*[7,22]. Sodium benzoate and phenylacetic acid are excreted in urine after conjugation with the nitrogen-containing compounds glycine and glutamine respectively. A patient with partial CPS deficiency was given 6 g sodium benzoate, and then 6 g of phenylacetate. On each occasion, urinary nitrogen excretion increased by approximately 50 per cent and blood ammonia levels (which were within the normal range whilst the patient was receiving a low-protein diet) and glutamine levels fell. Four patients in coma due to ammonia intoxication were given 200 - 350 mg per kg of sodium benzoate intravenously and again blood ammonia levels fell, this time from 100 times normal to near normal.

It seems that with a combination of therapeutic strategies long-term treatment of the severe forms of urea cycle disorder is becoming possible[23]. Nevertheless, many patients remain at risk of sudden onset of hyperammonaemic coma, due to infection or dietary indiscretion, and the incidence of sudden death remains high. Prompt treatment of infections and avoidance of catabolism by ensuring an adequate energy intake at all times is an important element of long-term treatment[24]. Peritoneal or haemodialysis, in skilled hands, effectively removes excess ammonia when other measures fail but these procedures need to be started before irreversible brain damage has occurred. The demands which these advances in treatment make on medical services and the families involved are great, and much more information is needed about long-term survival, intellectual progress and social adjustment before the real impact of these developments can be assessed.

1. Shih, E. in *The Metabolic Basis of Inherited Disease* (eds Stanbury, J.B., Wyngaarden, J. B. & Fredrickson, D. S.) 362 (McGraw Hill, New York , 1978).
2. Levin, B. & Russell, A. *Am. J. Dis. Child* **113**, 142 (1967).
3. Snyderman, S. E. *et al. Pediatrics* **56**, 65 (1976).
4. Walser, M. *et al. Clin.Sci.molec.Med.* **53**, 173 (1977).
5. Batshaw, M. L., Brusilow, S. & Walser, M. *Pediatrics* **58**, 227 (1976).
6. Brusilow, S. W. & Batshaw, M. L. *Lancet* **i**, 134 (1979).
7. Brusilow, S. W., Valle, D. L. & Batshaw, M. L. *Lancet* **ii**, 452 (1979).
8. Batshaw, M. L. & Brusilow, S. W. *J. Pediat.* **97**, 893 (1980).
9. Wiegard, C. *et al. J. Pediat.* **96**, 142 (1980).
10. Snyderman, S. E. *et al. J. Pediat.* **95**, 61 (1979).
11. Visek, W. J. *Nutr. Rev.* **37**, 273 (1979).
12. Rajantie, F., Simell, O. & Perheentupa, J. *Lancet* **i**, 1219 (1980).
13. Oyanagi, K. *et al. J. Pediat.* **94**, 255 (1979).
14. Valle, D., Walser, M. & Brusilow, S. W. *J.clin.Invest.* **65**, 371 (1980).
15. Hommes, F. A. *et al. Rep. int. Symp. on Metabolic Disease*, Interlaken (1980).
16. Collins, F. S. *et al. J. Pediat.* **96**, 429 (1980).
17. Bachmann, C. & Colombo, J. P. *Eur. J. Pediat.* **134**, 109 (1980).
18. Stewart, P. M. & Walser, M. *J.biol. Chem.* **255**, 5270 (1980).
19. Close, J. H. *New Engl. J. Med.* **290**, 663 (1974).
20. McReynolds, J. W. *et al. J. Pediat.* **93** 421 (1978).
21. Thoene, S. *et al. J. Pediat.* **90**, 213 (1977).
22. Brusilow, S., Tinker, J. & Batshaw, M. L. *Lancet* **ii**, 432 (1979).
23. Brusilow, S., Thomas, G. H. & Brusilow, S. W. *Pap. at Meet. on Inborn Errors of Metabolism*, Florida (1980).
24. Smith, I. & Francis, D. E. M. in *Textbook of Paediatric Nutrition* (eds McLaren, D. & Burman, D.) 273 (Churchill Livingstone, 1976).

Late-glacial climatic changes

from Peter Moore

CONSIDERABLE interest and controversy surround the climatic events at the close of the last glaciation. In north-west Europe, the warming which began at around 13,500 BP was interrupted by one major cold phase (at about 11,000–10,000 BP) and one less severe event (at about 12,000–11,800 BP) (dates according to Mangerud *et al. Boreas* **3**; 109, 1974). Both of these cold episodes are clearly expressed in the physical nature of lake sediments and in the contained pollen assemblages of sites in mainland north-west Europe, such as Denmark; but the earlier cold phase is not well marked either sedimentologically or palynologically in British sites, especially western ones. This has led Pennington (*Boreas* **4**; 157, 1975) to define a single 'Windermere' interstadial for the British Isles, effectively encompassing the time covered by the two continental interstadials.

Moving further east and south across Europe, one finds that the evidence for this complex of climatic oscillations is weaker, and the amplitude of the climatic fluctuations is evidently dampened. Watts (in *Studies in the Late-Glacial of North-West Europe* (eds J.J. Lowe, J.M. Gray & J.E. Robinson) Pergamon Press, Oxford; 1980) claims that there is no Alpine site which clearly supports a double oscillation of climate in the late-glacial and that even the major cold phase (the 'Younger Dryas' of northern Europe) may be difficult to define, though it is often marked by an opening of *Pinus cembra* and *Larix* forest and an invasion of *Juniperus* and *Artemisia*. Soil erosion at the time may also be reflected in an increased input of minerogenic sediments into lake basins. Data from the Mediterranean area and south-west Europe for the late-glacial period are still rather scanty and open to a variety of interpretations.

On the other side of the Atlantic, there have been pollen diagrams from North America which suggest a late-glacial (Wisconsin) climatic oscillation. Perhaps the best known of these is that of Schwegger (*Ecology* **50**; 859, 1969) in which a *Picea* and *Fraxinus*-dominated early stage of forest development is interrupted by a period in which the arboreal/non-arboreal pollen ratio falls and *Artemisia* increase. The warm phase, dated at 11,640 BP, has been termed the 'Two Creeks Interstadial' and geomorphological evidence shows that there were minor extensions and contractions of the ice sheet at the closing stages of the Wisconsin glaciation (Wright *Geol. Soc. Am. Mem.* **136**; 251, 1973). Some of the most detailed pollen data for this period have come from the work of Birks at Wolf Creek in Minnesota (*Ecol. Monogr.* **46**; 395, 1976). He has demonstrated the existence, between 20,500 and 14,700 BP, of tundra vegetation which was invaded by dwarf shrub with willow and alder, and then at about 13,600 BP, *Picea* became the dominant pollen type. Between 13,600 and 10,000 BP there was no evidence for substantial change in the vegetation, but invasion by *Pinus* subsequently followed. So, the climatic changes which resulted in movement of the residual ice front over that late-Wisconsin time period in Minnesota did not cause marked fluctuations in the vegetation around Wolf Creek. Birks considers it likely that the climatic thresholds necessary to cause such changes were not crossed. The inertia and inherent stability of the spruce forest evidently withstood the climatic oscillations, though their effect seems to have been felt more evidently in Schweger's site at Iola Bog, Wisconsin.

In the other direction, a pollen diagram from the Peking district of China (Kong and Du *Acta bot. sinica* **22**; 330, 1980), covering the period of about 30,000–10,000 BP, permits an examination of late-glacial events in central Asia. In the basal sediments (older than 22,700 BP) *Artemisia,* Chenopodiaceae and Gramineae dominate, indicating a cold, arid steppe-tundra. At about 13,000 BP *Abies, Picea* and *Pinus* expand and Artemisia contracts, as gymnosperm-dominated forests invaded the Peking plain. By 11,850 *Tilia* (a taxon demanding warmth) begins to expand, and the authors suggest this point as the base of the 'Holocene'. But before a radiocarbon date of 10,750 BP, there are two marked peaks in the pollen of *Artemisia,* the first being accompanied by *Thalictum* and the second by Chenopodiaceae. It is tempting to interpret these as times of reversal in the trend towards warmer climate, permitting once again the temporary extension of the steppe-tundra herbs.

The late-glacial climatic oscillation is not a European anomaly, as was once suggested by Mercer (*Arctic & Alpine Res.* **1**; 227, 1969), though it does seem to be both more marked and more consistently recorded in north-west Europe than elsewhere. This may, in part, be due to the current paucity of information from many areas of southern Europe and Asia. Attempts to correlate warm and cold episodes over wide geographical areas are frustrated by the complexity of sequences, local variations in the record (perhaps a consequence of the threshold problem) and, not least, by the inherent errors in the radiocarbon dating of sediments.

Peter Moore is in the Department of Plant Sciences, University of London King's College.

0028-0836/81/230380-01$01.00

REVIEW ARTICLE

The evolution of metabolic cycles

Jack E. Baldwin* & **Hans Krebs**†

*Dyson Perrins Laboratory (Organic Chemistry), South Parks Road, Oxford OX1 3QY, UK
†Metabolic Research Laboratory, Nuffield Department of Clinical Medicine, Radcliffe Infirmary, Oxford OX2 6HE, UK

The function of the citric acid cycle is to convert efficiently the energy released during the combustion of acetate into the energy stored in the pyrophosphate bonds of ATP. The cycle is almost twice as efficient as feasible alternatives of acetate combustion, such as a direct pathway via glycollate, glyoxylate, formaldehyde and formate. The reason is that the first stage of acetate degradation cannot be a dehydrogenation; it must be an oxygenation by molecular oxygen. Thus the cycle must have evolved because in a competitive environment the chances of survival are greatest if resources are optimal. Analogous considerations apply to the evolution of other metabolic cycles.

WHY have metabolic cycles, such as the citric acid cycle, arisen in the course of evolution? The starting point of the explanation presented here is a basic principle of evolution by natural selection which states that in a competitive environment, the chances of survival are greatest if optimal use is made of resources; or, as Orgel and Crick[1] have recently formulated it, replacing survival of the individual by survival of the species: the theory of natural selection in its more general formulation deals with the competition between replicating entities. In such a competition, the more efficient replicators increase in numbers at the expense of their less efficient competitors. After a sufficient time, only the most efficient replicators survive.

Definition of metabolic cycles

This discussion is limited, in the first instance, to 'metabolic cycles', defined as processes in which an overall chemical change is brought about by a cyclic reaction sequence. The citric acid cycle is a prototype: the overall change is the complete combustion of acetate, achieved by the cyclic formation of di- and tricarboxylic acids. Other types of cyclic processes exist where no overall change occurs, such as the storage cycles of glycogen or of triglyceride. In times of plenty, glucose is converted to glycogen, and in times of need, the glycogen is reconverted into glucose. Similarly, triglyceride is deposited when there is an excess source of energy and it is degraded by hydrolysis when more energy is needed than is supplied by the food.

The evolution of the citric acid cycle

The evolutionary axiom defined above leads to the question: does the citric acid cycle possess advantages over other pathways of acetate oxidation? It is the function of this cycle to provide energy in a form that can be utilized by living cells, that is, in the form of ATP. Hence the most advantageous mechanism of acetate oxidation is that which gives the highest yield of ATP. In evaluating the potential yield of ATP it has to be borne in mind that the bulk of ATP synthesis is not directly connected with the degradation of the foodstuff, but with the transport of hydrogen atoms (or electrons) from NADH to molecular oxygen. Thus it is a prerequisite of ATP synthesis that the primary process of degradation is a dehydrogenation.

The question has been considered whether a direct oxidation of acetate via glycollate, glyoxylate, formaldehyde and formate may be an alternative to the citric acid cycle[2]. An alternative possibility is an initial isomerization of acetic acid to glycolaldehyde, followed by the formation of glycollate. However, calculation of the equilibrium constant for this isomerization from standard heats of formation indicates that glycolaldehyde is about 76 kcal mol^{-1} less stable than acetic acid. Such a highly endothermic pathway must be ruled out. In the past no answer could be given to the question of why in living matter acetate is not 'directly' oxidized via glycollate. We now wish to offer an answer : the conversion of acetate or acetyl coenzyme A into glycollate cannot be brought about by a dehydrogenation. This reaction therefore cannot donate electrons to the mitochondrial electron transport chain. It can be oxidized only by molecular oxygen through monooxygenase ('hydroxylase' or 'mixed function oxidase'). This enzyme system, which is not located in the mitochondria but in the microsomal fraction, catalyses the insertion of one oxygen atom of molecular oxygen into the organic substrate while the second oxygen atom is reduced to water. The reduction requires a second substrate that donates the electrons; in most cases this is $\mathrm{NADPH_2}$. The general equation of an oxygenase reaction is therefore

$$\mathrm{R{\cdot}CH_3 + O_2 + NADPH_2 \rightarrow R{\cdot}CH_2OH + H_2O + NADP}$$

where $\mathrm{R{\cdot}CH_3}$ stands for any methyl group which, because of the nature of R, cannot be dehydrogenated. R can be —H, or —COOH, or, as in the case of toluene or xylene, a benzene ring.

Not only is an oxygenation of acetate wasteful because it cannot be coupled with ATP synthesis, but in addition, the consumption of one $\mathrm{NADPH_2}$ for the direct reduction of the second O atom of $\mathrm{O_2}$ robs the cell of the potential synthesis of three ATP molecules. The quantitative aspects of the waste are as follows: when free acetate is oxidized through the citric acid cycle there is a net gain of 10 pyrophosphate bonds of ATP. Twelve are formed by the reactions of the cycle but two pyrophosphate bonds are used for the conversion of acetate into acetyl coenzyme A.

When carbohydrate is the precursor, acetate arises in the form of acetyl coenzyme A directly from pyruvate. When long-chain fatty acids are the precursor of acetate, two pyrophosphate bonds are needed to form acyl-coenzyme A. Acetyl coenzyme A then arises by β-oxidation without any further expenditure of ATP. Hence, for a $\mathrm{C_{18}}$ saturated fatty acid, on average 2/9 pyrophosphate bonds are spent before acetate enters the cycle, and the ATP yield of the cycle is therefore $12 - 0.22$ or 11.78 pyrophosphate bonds for one turn of the cycle.

Were acetate 'directly' degraded via glycollate, then six ATPs would be prevented by the oxygenase reaction (with its need for $\mathrm{NADPH_2}$) from being formed. If all three stages of dehydrogenation from glycollate to $\mathrm{CO_2}$ were brought about by NAD- or NADP-linked dehydrogenases, then no more than six ATPs could be formed because one $\mathrm{NADH_2}$ is needed in the oxygenase reaction. This is a maximum value. The only glycollate-oxidizing enzyme (EC1.3.1.) known is a flavoprotein and therefore can at most yield two ATPs.

This concept does not only explain why the direct oxidation of acetate cannot compete efficiently with another pathway; it also offers an explanation of why this other pathway must be a cyclic one. The principle is this: when a metabolic sequence of reactions cannot occur efficiently by a direct route but necessitates a primary attachment of a substance to another molecule of low molecular weight (acetate to oxaloacetate), then the 'other molecule of low molecular weight' must be regenerated when the metabolic process of the substrate (acetate) has been

0028-0836/81/220381—02$01.00

completed. This is necessary, because were it not so, the other molecule of low molecular weight would have to be produced afresh. This would be wasteful, if not impossible: if the other molecule of low molecular weight were not regenerated it would yield a by-product in stoichiometric proportions, which would be nonsensical: an adult man on a diet of 2,000 kcal would produce 800 g oxaloacetate in 24 h. Thus a cyclic process is to be postulated as the only rational and economic mechanism for the organization of certain metabolic processes.

Optimal use of resources

The prediction that a cyclic pathway of acetate degradation is more efficient than the linear direct route of acetate oxidation does not postulate the specific details of the citric acid cycle. That acetate happens to react as the thioester of coenzyme A and to condense with oxaloacetate may be explained by the well documented assumption that the pathway leading to citrate and thence to α-oxoglutarate had evolved long before O_2 appeared in the atmosphere, in connection with the synthesis of glutamate. Coenzyme A is assumed to have arisen on the surface of the earth very early, possibly before 'life', together with the chemically closely related nucleic acids and other coenzymes[3-5].

Thus in the evolution of the citric acid cycle use was made of mechanisms already existing in connection with other functions. It is, indeed, a general principle of evolution that multiple use is made of given resources. Amino acids have many other functions apart from being building blocks of proteins. For example glycine is a precursor of creatine, bile acids, porphyrins, glutathione and hippurate; serine is a precursor of ethanolamine and choline, that is of phospholipids, and of acetylcholine. These multiple uses represent specific cases of the principle stated earlier concerning optimal use of resources.

Advantage of additional control points

Apart from being more efficient in making optimal use of the energy released by the combustion of acetate, the citric acid cycle has the advantage over a linear reaction sequence because it offers additional possibilities for metabolic control.

The cycle is regulated at several stages[6,7], one being the availability of oxaloacetate, the first step of the cycle. As oxaloacetate reacts like a catalyst, a cyclic organization has the advantage that the faster the rate of the cycle, the more rapid is the regeneration of the catalyst. Thus oxaloacetate is not only a catalyst of the ordinary type promoting a chemical process, but also provides control of a feedback type. Furthermore, the concentration of oxaloacetate can be adjusted by the redox state of the NAD-couple and by the tendency of aspartate aminotransferase to establish an equilibrium state of its substrates.

Other metabolic cycles

Some metabolic cycles are precisely analogous to the citric acid cycle in that the substrate, before it can react, has to combine with another substance of low molecular weight, and this other substance must then be regenerated. Examples are the ornithine cycle of urea synthesis, the Calvin cycle of photosynthesis and the glyoxylate cycle which brings about the synthesis of succinate from molecules of acetate. In the ornithine cycle, the catalyst also provides facilities for introducing a supply of energy in the form of ATP, needed for endothermic synthesis of urea.

The pentose phosphate cycle

Somewhat different in detail, but the same in principle, is the optimal efficiency achieved by the pentose phosphate cycle. The function of the pentose phosphate cycle is to generate $NADPH_2$ for reductive biosyntheses in the cytosol, above all for the synthesis of fat from carbohydrate. $NADPH_2$ is formed by glucose-6-phosphate dehydrogenase plus 6-phosphogluconate dehydrogenase. Thus one molecule of glucose-6-phosphate forms two molecules of $NADPH_2$, with ribulose-5-phosphate as a by-product. A complex series of reactions then reconverts this by-product to glucose-6-phosphate in such a way that six molecules of glucose-6-phosphate entering the cycle are converted to five molecules of glucose-6-phosphate while one undergoes complete degradation to 6 CO_2 and 12 $NADPH_2$. Thus NADP acts in place of O_2 as hydrogen acceptor in this degradation of glucose. The free energy changes in this pathway are slight; hence the pentose phosphate cycle disposes effectively and economically of an unwanted by-product.

The evolution of other metabolic pathways

The question of why certain metabolic pathways have evolved can be profitably raised also about non-cyclic mechanisms. What is the advantage of the pathway of propionate utilization via propionyl coenzyme A, methyl malonate coenzyme A and succinyl coenzyme A, as opposed to primary dehydrogenation to acrylate? What are the advantages of fatty acid synthesis via condensation of acetyl coezyme A with acyl coenzyme A? Why are ω-oxidation and the desaturation of saturated long-chain fatty acids brought about by monooxygenase, rather than dehydrogenation? A discussion of these question is beyond the scope of this article.

1. Orgel, L. E. & Crick, F. H. C. *Nature* **284,** 604–607 (1980).
2. Krebs. H. A. *Enzymologia* **12,** 88–100 (1947).
3. Oparin, A. I. *The Origin of Life on the Earth* 3rd edn (Oliver & Boyd, London, 1957).
4. Handler, P. in *Proc. 5th int. Congr. Biochem.* (ed. Oparin, A. I.) 149–157 (Macmillan, New York, 1961).
5. King, G. A. M. *Bio-systems* **13,** 23–45 (1980).
6. Bowman, R. H. *J. biol. Chem.* **241,** 3041–3048. (1966).
7. Randle, P. J., England, P. J. & Denton, R. M. *Biochem. J.* **117,** 677–695 (1970).

ARTICLES

Microwave OH maser emission in the circumstellar envelopes of late-type stars

P. F. Bowers[*,†], K. J. Johnston[†] & J. H. Spencer[†]

[*]Astronomy Program, University of Maryland, College Park, Maryland 20742, USA

[†]E. O. Hulbert Center for Space Research, Naval Research Laboratory, Washington DC 20375, USA

The total angular extents and absolute positions (± 1 arc s) of OH maser clouds surrounding 20 red giant and supergiant stars have been measured using the Very Large Array in a spectral line mode. Our results are consistent with a symmetric, expanding circumstellar shell model for many of the sources, but about half of the sources have more complex structure at the arc second level. A correlation is found between the sizes of the masing regions at 1,612 MHz, the expansion velocities of the circumstellar shells, and the estimated mass loss rates. For one star (U Orionis) initiation of extensive mass loss has occurred very recently and dramatic changes in its OH spectrum may represent the 'turn on' stage of OH masers in evolved stars.

0028-0836/81/220382—04$01.00

MICROWAVE OH maser emission at 1,612 MHz is commonly associated with long-period, variable, oxygen-rich red giant and supergiant stars which are undergoing extensive mass loss. The OH spectra characteristically have a double-peaked velocity structure in which the two strongest spectral features have velocity separations ranging from ~ 5 km s^{-1} to 50 km s^{-1}. Evidence that the stellar velocity is the mean velocity of the two OH peaks and that the OH masers can occur at large distances from the star[1–4] make an expanding, circumstellar shell model particularly attractive for these evolved stars. In this model the low-velocity (blueshifted) emission arises from the near side of the shell (as viewed from the Earth), whereas the high-velocity (redshifted) emission arises from the far side. The terminal expansion velocity V_e of the gas is one half the total velocity separation of the peaks.

Most of our direct knowledge of the structural characteristics of the masing regions associated with these stars is derived from very long baseline interferometric (VLBI) measurements. Unfortunately, such measurements generally provide only lower limits for the extent of the masing regions[3]. In addition, most VLBI maps of the spatial distribution of the maser components for these stars indicate that the distribution is much more complex than is predicted by the simple shell model[5]. Recent results obtained by Booth *et al.*[6] with relatively short VLBI baselines (~ 100 km) for the source OH127.8 − 0.0 are a dramatic exception, however, and demonstrate that at least for this star the expanding shell model is correct.

To obtain a better understanding of the overall size and structure of the OH masing regions for a significant number of these stars, we have used the National Radio Astronomy Observatory Very Large Array radio telescope near Socorro, New Mexico to obtain lower spatial resolution (~ 1.2 arc s synthesized half-power beamwidth at 18 cm) than is available with VLBI techniques. All observations were made during 22–23 November 1980 with a spectral line mode (frequency resolution = 6.1 kHz after Hanning-smoothing), nine antennas, and right circular polarization. Data from 20 stellar maser sources were obtained with a total integration time of ~ 1 h per source. Each maser source was observed at 3 or 4 hour angles, allowing a distribution of observations in the *uv* plane. Calibrators were observed every 30 min.

We present here only initial results obtained by separately integrating over the total velocity extents of the low-velocity and of the high-velocity emission complexes of each source. Because of the large number of sources which comprise our data, we will present an analysis of the extent and relative positions of individual spectral features and a detailed comparison with shell models elsewhere.

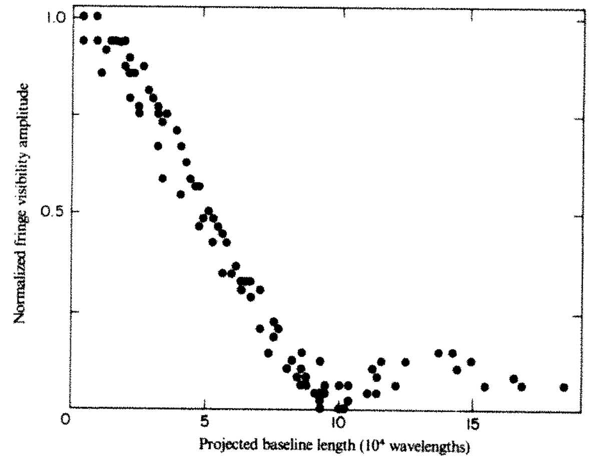


Fig. 1 Normalized fringe visibility as a function of projected baseline length for the integrated high velocity complex (+52.5–+55.4 km s^{-1}) of R Aquilae. All data for this and subsequent figures were obtained at a rest frequency of 1,612.231 MHz during epoch 1980.9.

Positions of the maser clouds

Table 1 lists for each source the peak position of the integrated flux density precessed to epoch 1950, the corresponding velocity range for the position determination, the circumstellar expansion velocity derived from total power OH spectra, the estimated distance, the total angular size of the masing region, the corresponding linear size, the estimated mass loss rate $\dot{M}$, and the corresponding reference for $\dot{M}$.

In all cases the velocity range chosen for the position determination was dominated by a relatively strong, narrow spectral component and the resulting maps showed a fairly simple structure with a well-defined maximum intensity. We estimate that the positions listed in Table 1 are accurate to ±1.0 arc s. The relative positions between the high- and low-velocity components are accurate to ±0.3 arc s.

For stars where both the low and high integrated velocity maps showed simple structure, positional offsets between the two complexes were $\lesssim 0.3$ arc s. To within our uncertainties, this

Table 1 Positions, shell sizes, and mass loss rates for OH/IR stars

Star	Position of emission peak α(1950)	δ(1950) (km s^{-1})	Velocity range (km s^{-1})	V_e (kpc)	Distance (arc s)	Angular size ($\times 10^3$AU)	Linear size	$\dot{M}$ ($\times 10^{-6} M_\odot$yr^{-1})	Ref.
IRC+10011	01 h03 min 48.08s	+12°19′51.2″	−10.9–−7.4	18	0.5	2.6×2.2*	1.3	17	7
OH127.8−0.0	01 30 27.63	+62 11 31.2	−67.5–−61.2	11	3.3	2.2	7.3	32†	
OH138.0+7.3	03 20 41.48	+65 21 32.8	−48.8–−45.4	10	2.4	1.4	3.4	12†	
OH141.7+3.5	03 29 23.66	+60 10 04.4	−71.0–−67.6	12	3.7	0.9	3.3	17†	
IRC+50137	05 07 19.68	+52 48 53.7	+18.3–+21.1	17	0.8	2.4	1.9	11	7
U Ori	05 52 50.93	+20 10 06.1	−43.5–−40.1	3	0.2	≲1	≲0.2	1.2	11
IRC+40156	06 29 45.03	+40 45 08.2	−29.0–−26.8	12	1.4	1.1	1.5	8†	
VY CMa	07 20 54.72	−25 40 12.5	−14.6–+8.1	28	1.5	2.5×1.4*	3.8	165,600	7,13
IRC−20197	09 42 56.55	−21 47 54.2	+26.5–+28.8	12	0.7	1.0	0.7	4	7
WX Ser	15 25 31.98	+19 44 13.0	−1.9–+0.9	7	1.0	0.8	0.8	2,8	7,11
OH17.7−2.0	18 27 39.77	−14 31 03.9	+48.0–+53.6	12	5.4	3.1×1.4*	16.8	88†	
OH26.5+0.6	18 34 52.47	−05 26 37.2	+29.3–+42.9	14	2.2	4.7×2.2*	10.4	400,70,200	10,12,13
OH25.1−0.3	18 35 33.36	−07 12 35.2	+129.0–+131.8	13	9.1	1.4	12.8	78†	
OH32.8−0.3	18 49 48.15	−00 17 53.2	+44.1–+47.5	16	4.3	3.2×2.1*	13.8	20,50	10,12
OH39.7+1.5	18 56 03.88	+06 38 49.8	+30.2–+38.1	16	1.4	2.8×1.8*	3.9	20,50	10,13
OH39.9−0.0	19 01 42.90	+06 08 44.2	+159.3–+165.0	15	7.7	1.6×1.0*	12.3	100	10
R Aql	19 03 57.67	+08 09 07.7	+52.5–+55.4	6	0.3	3.1	0.9	0.8	11
IRC+10420	19 24 26.74	+11 15 11.2	+40.4–+63.1	29	3.4	3.3×2.0*	11.2	340†	
RR Aql	19 55 00.30	−02 01 17.5	+31.4–+34.8	5	0.4	1.0	0.4	0.4†	
NML Cyg	20 44 33.84	+39 55 57.2	−32.8–+1.3	23	0.5	3.8×3.2*	1.9	36,13	7,11

* Indicates complex (multiple-component) structure.
† Values determined from equation (1).

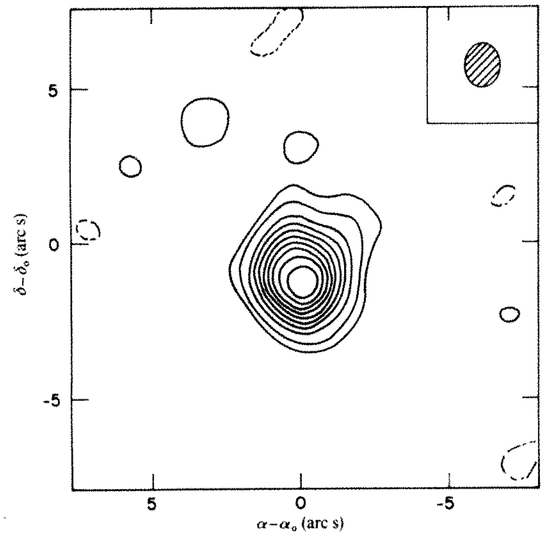


Fig. 2 Angular intensity distribution as a function of position for the integrated high-velocity complex of R Aquilae. Contour levels are in units of −5%, 5%, 10%, 20%, - - - , 90% of the peak integrated flux density. The FWHM synthesized beamwidth is shown in the north-west quadrant of the diagram. The map centre is α_0 (1950) = 19 h 03 min 57.68 s, δ_0 (1950) = +08°09′09.1″.

result supports a symmetric, expanding shell model for most of the stars. (Our integrated velocity data do not allow us to distinguish between spherical symmetry and other symmetric, expanding geometries.) As this model predicts that the position of maximum maser gain occurs along the line of sight to the star, it is likely in these simple cases that our positions are equivalent to the stellar positions. This view is supported for the bright stars U Ori and R Aql as their positions listed in Table 1 agree with the SAO optical positions to within 1 arc s after correction for the effects of proper motion.

Structure of the maser clouds

Except possibly for U Ori which may be unresolved, all our sources display structure on an arc second scale. About one half of them need to be modelled by more than one spatial component and are classified as complex in Table 1. The large number of sources with simple (single-component) structure may be due to the dominance of the strongest spectral feature in each integrated emission complex.

To illustrate our data for a source with relatively simple structure, we show the visibility function (Fig. 1) and the intensity distribution (Fig. 2) for the high-velocity complex of the Mira variable R Aquilae. The curve in Fig. 1 is a classic visibility function for a circular-symmetric source with a uniform brightness distribution. This result is especially significant as most previous size estimates of stellar maser components assume gaussian brightness distributions. The derived size from Fig. 1 agrees well with the total extent of 3.1 arc s obtained from the map in Fig. 2 after correction for source broadening due to the beam. The slight asymmetry in the emission distribution in Fig. 2 may be an instrumental effect due to inadequate phase calibration for this source. Note, however, that the centroids of the much weaker low velocity complex (not shown) and the high-velocity complex agree in position to within 0.3 arc s, which supports a simple shell model for this source.

As an example of a source with complex structure, Fig. 3 shows the angular distribution of the integrated emission of the low-velocity and high-velocity emission complexes for the M supergiant star NML Cygni. Figure 3 indicates significant differences between the structure of the integrated emission from the two velocity complexes. The differences do not support a simple expanding shell model where the two emission complexes should appear to be spatially coincident at the stellar position. On the other hand, our maps are in very good qualitative agreement with VLBI maps[5] which show the relative positions of 1,612-MHz maser components towards this star. Points of similarity include: (1) the central location of the low-velocity emission in the overall masing cloud; (2) positionally separate concentrations of low- and high-velocity emission; and (3) an elongated distribution of the high-velocity emission with a position angle of ~135° and a total angular extent of several arc seconds.

Angular sizes of the maser clouds are listed in Table 1. All values were obtained from measurements of the total angular extent of the integrated emission to the lowest significant contour level (typically ~10%) and have been corrected for source broadening due to the synthesized beam. In cases where the sizes of the low- and high-velocity complexes are different, we list the larger size. Our size estimates may sometimes be too small because of the limited dynamic range of the maps. For sources with simple structure, however, sizes determined from the maps are in relatively good agreement with sizes determined from the visibility functions if a uniform disk brightness distribution is assumed. We cannot, of course, rule out the possibility of low-level emission (≤10% of the total integrated emission) with much greater scale sizes.

Relation between OH structure and stellar mass loss

We have recently assessed[3] the available data pertaining to the structure of OH masing regions in late-type stars and suggested that the overall size of the 1,612-MHz masing region may be related to the stellar mass loss rate. That analysis was hindered, however, by the few OH/IR stars for which measurements of the mass loss rate and the shell size exist. The present data along with recent estimates of the mass loss rates

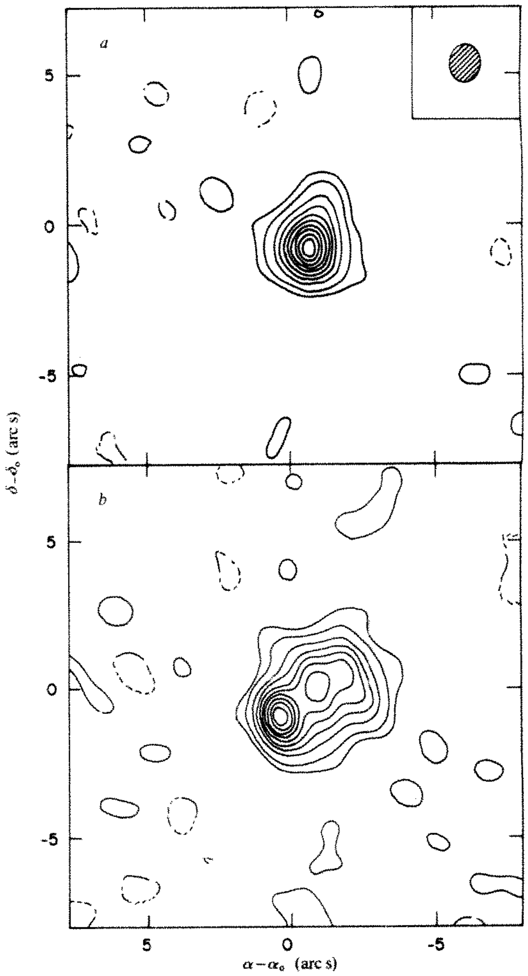


Fig. 3 Angular intensity distribution as a function of position for: *a*, the integrated low-velocity (−32.8 to +1.3 km s^{-1}); and *b*, the integrated high-velocity (+2.5 to +23.3 km s^{-1}) emission complexes of NML Cygni. Contour levels have the same relative percentages of the peak integrated flux as in Fig. 2. The map centre is α_0 (1950) = 20 h 44 min 33.90 s, δ_0 (1950) = +39°55′58.0″.

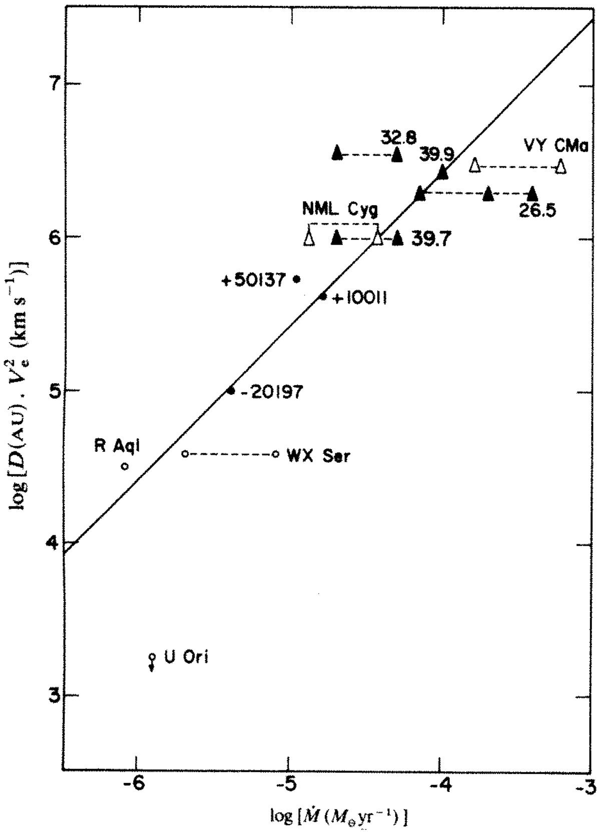


Fig. 4 Product of the total angular extent D(AU) of the 1,612-MHz masing region and the square of the circumstellar expansion velocity V_e (km s^{-1}) as a function of the stellar mass loss rate $\dot{M}$ ($M_\odot$ yr^{-1}). ○, Optically identified Mira variables; △, M supergiants. ●, Mira variables commonly designated as 'IRC' sources. ▲, Unidentified Type 2 OH/IR stars, designated as 'OH' stars in Table 1. The solid line corresponds to equation (1).

for several additional OH/IR stars now provide a much better data base to test our proposed relation.

To compare the sizes of the masing regions in these stars, we have converted the angular diameters to linear diameters using the distances listed in Table 1. Distances were obtained from ref. 7 for most of the optically identified or 'IRC' stars, from ref. 3 for the 'OH' stars, from ref. 8 for IRC + 10420, NML Cyg, and VY CMa, and a period–luminosity relationship[9] for U Ori. In general, the distance estimate (and thus the derived linear shell size) is probably only accurate to about a factor of 2.

Estimates of the stellar mass loss rates have been obtained for 12 stars in our sample from measurements of the depth of the 10-μm silicate absorption feature present in the IR spectra of many of these stars[10,11], from considerations of momentum conservation in the radiation-transfer process[12], or from model-fitting to the IR excess[7,13]. For stars observed by Hyland *et al.*[7], we have derived the total mass loss rate by adopting a canonical value of 100 for the mass ratio of gas to dust in the circumstellar shells. This value is consistent with values found by Werner *et al.*[12] for several OH/IR stars. If more than one estimate of the mass loss rate is available, we have listed all estimates to give an idea of the uncertainties of these derived quantities.

To investigate a possible relationship between the shell diameter D and the stellar mass loss rate $\dot{M}$, we consider a spherically expanding shell with density $n \propto \dot{M}/(r^2 V_e)$ at distance r from the star. Pumping models[14] indicate that the critical parameter for inversion of the OH rotational transitions is the optical depth τ_{OH}. If the maximum extent of the maser region is determined by the position at which τ_{OH} has a minimum value, we have $\tau_{min} \propto nr/V_e \propto \dot{M}/(DV_e^2)$. Assuming τ_{min} is constant from star to star, we then have $\dot{M} \propto DV_e^2$, where D and V_e are observables if the distance to the star is known.

The results plotted in Fig. 4 support such a correlation for the stars in Table 1 with values of $\dot{M}$ derived from IR data. A least-squares fit to all the data points in Fig. 4 except U Ori yields $DV_e^2 = 6.8 \times 10^9 \dot{M}^{(0.87 \pm 0.14)}$ with a linear correlation coefficient of 0.90. However, in view of the uncertainties in the data, we have drawn the solid line in Fig. 4 with a slope of 1, corresponding to

$$\log[D(\text{AU}) \cdot V_e^2\,(\text{km s}^{-1})] = 10.44 + \log \dot{M}\,(M_\odot\,\text{yr}^{-1}) \quad (1)$$

Equation (1) has been used to compute $\dot{M}$ for the remaining stars in Table 1. As values of $\dot{M}$ obtained from IR measurements depend on various critical assumptions concerning the properties of the dust, there may be considerable errors and the scale of the values for $\dot{M}$ may be uncertain. Nevertheless, the correlation in Fig. 4 suggests that the relative values of $\dot{M}$ for these stars may be approximately correct to an order of magnitude.

The correlation in Fig. 4 also suggests that the expansion time scales of the shells are greater than $D/(2V_e)$ for most of these stars. A prominent exception is U Ori whose size ($\leq$200 AU) is the smallest of all the stars we observed (Table 1). This star is particularly noteworthy because of recent dramatic changes in its OH spectrum (from a Type 1 or main-line emitter to a Type 2 or 1,612-MHz emitter)[15,16] and because of changes in the structure of the OH masing region wherein the apparent sizes of the maser components have increased 10-fold in 3 yr (ref. 17). Such dynamic events have thus far been observed only for this star.

We suggest that initiation of extensive mass loss has occurred fairly recently ($\leq$150 yr) for this star. The phenomena observed for the OH masers from U Ori may be manifestations of the physical conditions in the outer regions of a young, expanding envelope whose density and temperature (and thus maser pump conditions) are rapidly changing. The conversion of the OH spectrum from Type 1 to Type 2 is consistent with this model based on theoretical suggestions that the main-line inversion region is inside the 1,612-MHz inversion region[18]. The dynamic phenomena observed for this star thus may be representative of the 'turn on' stage of masers associated with late-type stars.

Conclusions

Absolute mean positions (±1.0 arc s) have been determined for OH maser clouds surrounding 20 red giants and supergiants. In most cases the derived positions are probably the stellar positions.

The structure of the masing regions around many of the stars supports a simple, symmetric, expanding shell model, but there are numerous cases where the structure is significantly more complex.

Diameters of the masing regions vary from $\sim 10^2$ AU to $> 10^4$ AU and are proportional to $\dot{M}/V_e^2$.

Because of the extremely small shell size observed for one star (U Ori) and because of the recent dramatic changes in its OH spectrum, we suggest that the star is surrounded by a young ($\leq$150 yr), expanding envelope whose OH masers have only recently been turned on.

We thank Drs R. D. Gehrz, S. G. Kleinmann and C. M. Telesco for results before publication and Dr P. R. Schwartz for helpful comments. The National Radio Astronomy Observatory is operated by Associated Universities, Incorporated, under contract with the NSF.

Received 19 January; accepted 10 April 1981.

1. Dickinson, D. F., Reid, M. J., Morris, M. & Redman, R. *Astrophys. J. Lett.* **220**, L113–L116 (1978).
2. Reid, M. J., Muhleman, D. O., Moran, J. M., Johnston, K. J. & Schwartz, P. R. *Astrophys. J.* **214**, 60–77 (1977).
3. Bowers, P. F., Reid, M. J., Johnston, K. J., Spencer, J. H. & Moran, J. M. *Astrophys. J.* **242**, 1088–1101 (1980).
4. Jewell, P. R., Webber, J. C. & Snyder, L. E. *Astrophys. J. Lett.* **242**, L29–L31 (1980).
5. Benson, J. M. & Mutel, R. L. *Astrophys. J.* **233**, 119–126 (1979).
6. Booth, R. S., Kus, A. J., Norris, R. P. & Porter, N. D. *Nature* **290**, 382 (1981).
7. Hyland, A. R., Becklin, E. E., Frogel, J. A. & Neugebauer, G. *Astr. Astrophys.* **16**, 204–219 (1972).
8. Mutel, R. L., Fix, J. D., Benson, J. M. & Webber, J. C. *Astrophys. J.* **228**, 771–779 (1979).
9. Clayton, M. L. & Feast, M. W. *Mon. Not. R. astr. Soc.* **146**, 411–421 (1969).
10. Mason, S., Gehrz, R. D., Kleinmann, S. G., Hackwell, J. & Grasdalen, G. (in preparation).
11. Gehrz, R. D. & Woolf, N. J. *Astrophys. J.* **165**, 285–294 (1971).
12. Werner, M. W. *et al. Astrophys. J.* **239**, 540 (1980).
13. Telesco, C. M., Kleinmann, S. G., Harper, D. A., Loewenstein, R. & Moseley, S. H. (in preparation).
14. Elitzur, M., Goldreich, P. & Scoville, N. *Astrophys. J.* **205**, 384–396 (1976).
15. Pataki, L. & Kolena, J. *Bull. Am. astr. Soc.* **6**, 304 (1974).
16. Cimerman, M. *Astrophys. J. Lett.* **228**, L79–82 (1979).
17. Reid, M. J. *et al. Astrophys. J. Lett.* **227**, L89–92 (1979).
18. Elitzur, M. *Astr. Astrophys.* **62**, 305 (1978).

Molecular analysis of the *unc*-54 myosin heavy-chain gene of *Caenorhabditis elegans*

Alexander R. MacLeod, Jonathan Karn & Sydney Brenner

Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

The properties of a small internal deletion mutant, E675, have been exploited in the molecular cloning of the unc-*54 gene. This mutation uniquely identifies* unc-*54 sequences as molecules of altered length in E675 and provides a genetic and physical marker for the active gene, its messenger RNA and its protein product.*

THE *unc*-54 gene of *Caenorhabditis elegans* specifies a major myosin heavy chain present only in the body wall musculature of this small soil nematode[1-3]. All mutants in this gene are severely paralysed, and have disorganized muscle in which the numbers of thick filaments are reduced by 75–80% (refs 4, 5). However, these animals are able to feed and survive because the pharyngeal muscle is unaffected and continues to function normally.

A large collection of independently induced mutations in *unc*-54 is being analysed in this laboratory as part of a detailed study of this eukaryotic gene. Genetical techniques which allow selection of additional mutants, including deletions, have recently been developed (R. P. Anderson and S. Brenner, unpublished) and several suppressor genes which suppress specific alleles of *unc*-54 have been identified[6,7].

Many *unc*-54 alleles have been characterized by chemical cleavage of purified nematode myosin heavy chains[4], made possible by the abundance of the body wall muscle contractile proteins and the ease with which they can be isolated from homogenates of the animal. Strains carrying dominant or semi-dominant mutations were found to contain normal amounts of the *unc*-54 myosin heavy chain which were unable to assemble into functional thick filaments. Although most of these alleles seem to be missense mutations, in one particular semi-dominant mutant, E675, the *unc*-54 myosin heavy chain was reduced in molecular weight (MW) by an internal deletion of 100 amino acids near the C-terminus of the molecule[8].

Most recessive alleles of *unc*-54, such as *e*190, fail to produce stable protein products and cannot be studied by these methods. The present experiments use cloned complementary DNA and genomic DNA probes to define the structure of the *unc*-54 gene, and describe the assay of *unc*-54 messenger RNA by cell-free protein synthesis and hybridization. These methods allow detailed study of the recessive alleles of *unc*-54. This is of particular interest because we expect that this class of alleles includes nonsense mutations as well as mutations affecting mRNA synthesis and maturation.

Cell-free synthesis of the *unc*-54 myosin heavy chain

RNA was isolated from nematodes using a modification of the guanidinium chloride extraction procedure[9]. This yielded undegraded mRNA which supported high levels of protein synthesis in cell-free systems derived from both wheat-germ[10] and rabbit reticulocytes[11].

Total RNA from wild-type animals (N2) programmes the synthesis of three polypeptides of MWs 210,000, 205,000 and 200,000 which have the properties of myosin heavy chains. The putative myosin heavy chains synthesized in the wheat-germ cell-free system were examined by SDS-polyacrylamide gel electrophoresis (Fig. 1). Each of these proteins is quantitatively precipitated by antibody prepared against nematode myosin (purified by affinity chromatography on myosin–Sepharose) and each co-migrates with a heavy-chain component of nematode myosin. A similar result is observed when the reticulocyte cell-free system is programmed with nematode RNA, except that the 210,000-MW component is preferentially synthesized. We therefore worked primarily with the wheat-germ system, in

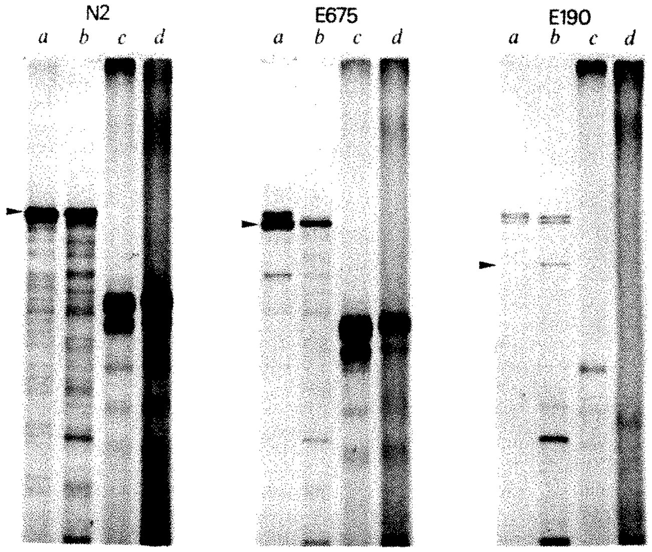


Fig. 1 Cell-free synthesis of nematode myosin heavy chains. RNA was isolated by guanidinium chloride extraction[9] of nematodes from worms grown in liquid medium[33]. The animals were disrupted by two passages through a French pressure cell at 12,000 p.s.i. Cuticular debris was removed by centrifugation at 30,000*g* for 15 min and RNA selectively precipitated at −20 °C for 12–15 h by addition of 0.5 vol 95% ethanol. The precipitated RNA was resuspended in 7 M urea, 0.3 M NaCl, 10 mM Tris-HCl *p*H 7.4, 1 mM EDTA and 0.2% SDS, and residual protein was removed by three successive phenol/chloroform/isoamyl alcohol extractions (25 : 24 : 1). The yield of RNA was usually 5 mg per g wet weight nematodes. Extracts of commercial raw wheat germ (Spillers Research) were prepared as described previously[10] but without preincubation. Protein synthesis was usually performed in a 50-μl reaction mixture containing 15 μl of gel-filtered wheat-germ S-30, 20 μg total nematode RNA, 1 mM ATP, 0.1 mM GTP, 10 mM creatine phosphate, 50 μg ml^{-1} creatine kinase (Boehringer), 20 mM HEPES (adjusted to *p*H 7.6 with KOH), 1 mM dithiothreitol (DTT), 96 mM KCl, 1.5 mM Mg-acetate, 0.8 mM spermidine, 200 μCi ml^{-1} ^{35}S-methionine (400–600 Ci mmol^{-1}, Amersham), 5 μM methionine and 40 μM each of the other 19 amino acids. Nematode RNA preparations frequently contained RNA aggregates which inhibited the synthesis of high molecular weight proteins. Incubation of the RNA in 10 mM Tris-HCl *p*H 7.4 and 0.1 mM EDTA at 65 °C for 5 min just before addition to the cell-free system eliminated this problem. Immunoreactive proteins were identified by incubating the wheat-germ reaction mixture with antiserum prepared against nematode myosin and purified by affinity chromatography on myosin–Sepharose. The antigen–antibody complexes were indirectly immunoprecipitated with formaldehyde-fixed *Staphylococcus aureus*[34] (MRE, Porton). To obtain sufficient material for analysis of the cyanylation reaction[8,13,14,35] the standard wheat-germ reaction mixture was increased to 1 ml. The myosin heavy chains were recovered by immunoprecipitation as described previously[34] except that the proteins which bound to the immunoabsorbent were solubilized in 6 M guanidinium chloride, 100 mM Tris-HCl *p*H 8.0 and 10 mM DTT at 37 °C for 1 h. Chemical cleavage was performed in the presence of 100 μg ml^{-1} of rabbit skeletal muscle myosin and the cyanylation fragments from the rod portion of the myosin molecule were analysed by SDS-polyacrylamide gel electrophoresis. ^{35}S-labelled myosin was extracted from nematodes grown on ^{35}S-labelled *E. coli*[2]. Gels[36] were prepared for fluorography[37] and exposed to pre-fogged film[38] at −80 °C. The autoradiographs depicted above show the myosin heavy chains of N2, E675 and E190 animals. In each panel: *a*, myosin heavy chains synthesized *in vivo*; *b*, myosin heavy chains synthesized by the cell-free system and purified by immunoprecipitation; *c*, cyanylation cleavage of myosin heavy chains synthesized *in vivo*; *d*, cyanylation cleavage of myosin heavy chains synthesized by the cell-free system. The arrows indicate the mobility of *unc*-54-specific cell-free translation products from each of the three strains.

0028-0836/81/220386—05$01.00

which the proportions of heavy-chain synthesis more accurately reflect the *in vivo* levels.

The major myosin heavy chain synthesized *in vitro* seems to be the product of translation of *unc*-54 mRNA. It co-migrates with the 210,000-MW *unc*-54 myosin heavy chain synthesized in wild-type animals, and shows altered patterns of synthesis when the cell-free system is programmed with RNA from *unc*-54 mutants. When RNA from E675 animals is used as a source of mRNA, this heavy-chain component is replaced by an equally abundant, but novel, polypeptide of MW 200,000, which now co-migrates with the major myosin heavy chain present in E675 animals. Our previous chemical cleavage studies of E675 myosin demonstrated that the truncated heavy chain arises from an internal deletion of 100 amino acids[8]. As an apparently identical deletion is observed in the E675 myosin heavy chain synthesized *in vitro*, it seems likely that the deletion is present in the *unc*-54 mRNA and is not due to post-translational modification.

The wild-type *unc*-54 heavy chain is also absent from the translation products of E190 RNA. However, a polypeptide of MW 170,000 can be detected among the products immunoprecipitated from the wheat-germ extract. This component is present in low yield and is found only in E190, suggesting that it may be a product of the *unc*-54 mRNA altered by the *e*190 mutation. We have failed to detect a polypeptide of similar molecular weight in the immunoprecipitated myosin fractions of E190 animals, probably because the shortened E190 heavy chain fails to assemble correctly into a stable myosin molecule and is therefore subjected to proteolytic degradation. In any case, the low level of *in vitro* synthesis of this product indicates that the *e*190 mutation causes defects in either the synthesis or survival of the *unc*-54 mRNA.

The other recessive null alleles assayed by cell-free translation also fail to synthesize full-length *unc*-54 myosin heavy chain[12]. Allele-specific peptides smaller than the wild-type *unc*-54 myosin heavy chain have been observed in immunoprecipitates of the translation products of E1092, E1201 and E1223 but not in E1108. The simplest explanation for this is that these peptides are each the products of the mutant *unc*-54 mRNAs. However, the low yields of these products have limited their further analysis.

Direct chemical evidence for the synthesis of the *unc*-54 myosin heavy chain in the cell-free system was obtained from an analysis of the large polypeptides produced by chemical cleavage of the heavy chains at cysteine residues by cyanylation[4,8,13,14]. The major cleavage products of the wild-type *unc*-54 myosin heavy chain are two fragments of MWs 150,000 and 140,000 (ref. 4). Cyanylation of the *unc*-54 myosin heavy chain from E675 produces an analogous pattern but the large peptides are reduced in molecular weight by 10,000 because of the deletion[8]. This characteristic cyanylation doublet is absent from cleaved E190 myosin heavy chains[4]. As shown in Fig. 1, cleavage of the myosin heavy chains synthesized in the wheat-germ extracts generates the correct cyanylation peptides. This demonstrates the correct placement of the internal cysteine residues and provides conclusive evidence that the *unc*-54 myosin heavy chain is synthesized *in vitro*.

The myosin heavy chains of MWs 205,000 and 200,000 seem to correspond to the previously identified pharyngeal and residual body wall myosin heavy chains, respectively[2–4]. We also detect a previously uncharacterized minor component of MW 210,000 in E675 and E190 animals. We assume that N2 animals also synthesize this product but that its presence is obscured by the *unc*-54 myosin heavy chain. This component is not synthesized *in vitro* in either the wheat-germ or rabbit reticulocyte cell-free systems and has not been characterized further.

Purification of the *unc*-54 mRNA

The preceding experiments demonstrate that cell-free translation of nematode mRNAs in a heterologous protein-synthesizing system provides a reliable assay for *unc*-54 mRNA activity. We applied this assay to monitor the purification of the myosin heavy-chain mRNAs. Because of their large size a considerable purification of these mRNAs can be achieved by successive sedimentation through neutral and denaturing[15] sucrose gradients. Figure 2 shows a typical purification of *unc*-54 mRNA.

Although a substantial purification of the *unc*-54 myosin heavy-chain mRNA has been achieved, this fraction still contains the mRNAs coding for minor myosin heavy chains as well as small amounts of mRNAs coding for unidentified proteins of lower molecular weight (Fig. 2*C*), and ribosomal RNA sequences which account for much of the nucleic acid mass. These stable RNA contaminants can, in principle, be removed by affinity chromatography of the *unc*-54 mRNA on oligo(dT)–cellulose[16] or poly(U)–Sepharose[17]. However, we have found that nematode myosin heavy-chain mRNAs, like other large mRNAs[18], are not retained efficiently on these columns, apparently due to steric hindrance rather than to the absence of a 3′-poly(A) tract.

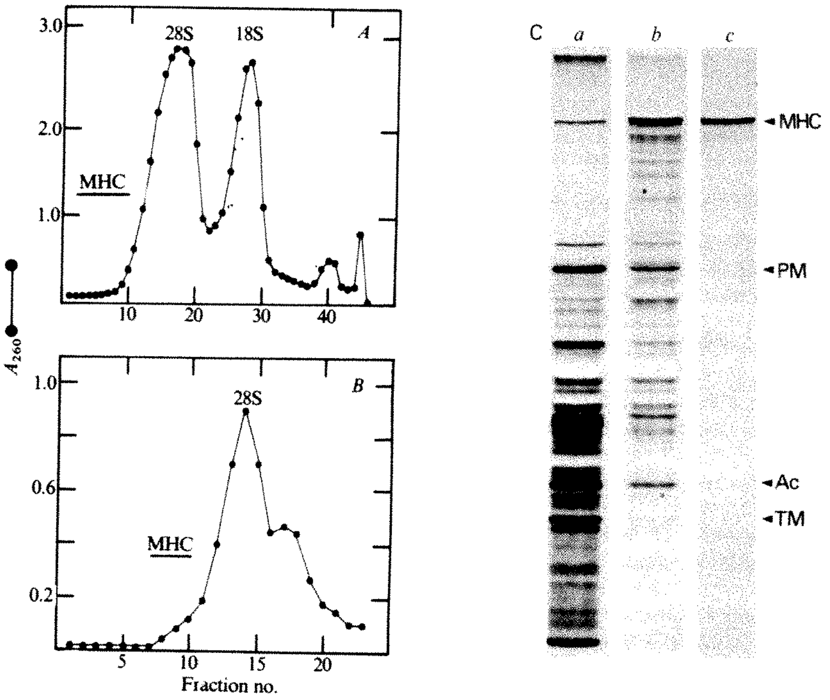


Fig. 2 Purification of the *unc*-54 mRNA. RNA (30 mg) from nematode larvae was dissolved in 10 ml of 10 mM Tris-HCl *p*H 7.5, 1 mM EDTA and 10% dimethyl sulphoxide, heated at 65 °C for 10 min and loaded on a 10–40% sucrose gradient in 100 mM NaCl, 10 mM Tris-HCl *p*H 7.5, 1 mM EDTA and 0.1% SDS generated in an MSE type XIV zonal rotor. After centrifugation at 40,000 r.p.m. for 15 h the gradient was unloaded and fractions containing myosin heavy-chain mRNA activity were identified by cell-free translation. These fractions were pooled, ethanol precipitated and redissolved in water at a concentration of 10 mg ml^{-1}. This was adjusted to 20 mM sodium borate *p*H 8.0 and 50 mM methylmercury hydroxide (Alfa)[15]. A 300-μl aliquot was loaded on a 13-ml 5–20% sucrose gradient in 100 mM NaCl, 10 mM Tris-HCl *p*H 7.5, 1 mM EDTA and 0.1% SDS. The gradient was centrifuged in the Beckman SW40 at 40,000 r.p.m. for 5 h. *A*, optical density profile of RNA fractionated in the zonal rotor. *B*, optical density profile of RNA refractionated after treatment with methylmercury hydroxide. *C*, assay of myosin heavy-chain mRNA purification by cell-free protein synthesis. *a*, Proteins synthesized in response to total larval RNA; *b*, proteins synthesized in response to the myosin mRNA purified by zonal gradient centrifugation, as shown in *A*: *c*, proteins synthesized in response to myosin mRNA repurified by the denaturing sucrose gradient, shown in *B*.

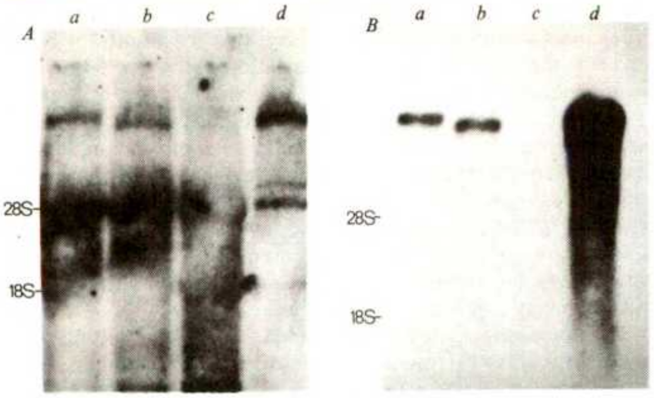


Fig. 3 Identification of plasmids containing sequences complementary to the *unc*-54 mRNA. cDNA was synthesized[19] in a 150-μl reaction containing 50 mM KCl, 50 mM Tris-HCl *p*H 8.3, 10 mM Mg-acetate, 2 mM DTT, 1 mM each of dGTP, dCTP, dTTP, 200 μM dATP and 50 μCi [d-^{32}P]dATP (Radiochemical Centre), 30 μg ml^{-1} oligo(dT), 5 μg enriched *unc*-54 mRNA and 40 units of AMV reverse transcriptase (gift of J. Beard). RNA was hydrolysed by incubation in 0.1 M NaOH and 1 mM EDTA[39] at 70 °C for 20 min, desalted[40] and the cDNA was made double stranded in a 100-μl reaction containing 100 mM KCl, 50 mM HEPES adjusted to *p*H 6.9 with KOH, 5 mM Mg-acetate, 1 mM DTT, 500 μM each dGTP, dCTP, dTTP, dATP and 50 μg ml^{-1} of DNA polymerase[20]. Single-stranded cDNA was digested with 8 units of S_1 nuclease[18] at 37 °C for 45 min in a 50-μl reaction mixture containing 200 mM NaCl, 50 mM Na-acetate *p*H 4.6 and 3 mM $ZnCl_2$. Fractions containing high molecular weight double-stranded cDNA eluting from a Sepharose 4B column were pooled and digested with 5 units *Mbo*I[21] at 37 °C for 30 min in 6 mM NaCl, 6 mM Tris-HCl *p*H 7.5, 6 mM Mg-acetate and 6 mM 2-mercaptoethanol. The *Mbo*I-digested DNA was ligated to *Bam*-digested pBR322[22] (previously treated with alkaline phosphatase[41]) in a 50-μl reaction containing 60 mM Tris-HCl *p*H 7.6, 8 mM Mg-acetate, 7 mM 2-mercaptoethanol, 0.25 μg vector DNA and 20 μg ml^{-1} T4 ligase. After incubation at 10 °C for 1 h, the reaction was diluted to 100 μl and mixed with 200 μl $CaCl_2$-treated cells[42]. Transformed cells harbouring plasmids containing inserts were selected as ampicillin-resistant, tetracycline-sensitive colonies. Plasmid DNA was prepared from colonies which hybridized to cDNA[43] by the caesium chloride–ethidium bromide centrifugation procedure[44]. These were labelled with [α-^{32}P]dATP by nick translation (see Fig. 6 legend) and tested for the presence of *unc*-54 sequences by hybridization to filters containing fractionated nematode RNA. RNA from N2, E675 and E190 animals and the partially purified *unc*-54 mRNA were separated in a 1.5% agarose gel in TBE buffer (90 mM Tris (base), 90 mM boric acid, 2.5 mM EDTA *p*H 8.3) containing 0.1% SDS. The RNA was transferred to diazotized paper[23] and hybridized to labelled probes. *A*, ^{32}P-cDNA prepared from enriched *unc*-54 mRNA probe. *B*, nick-translated pMbo7 DNA probe. In each panel: *a*, N2 RNA; *b*, E675 RNA; *c*, E190 RNA; *d*, partially purified *unc*-54 mRNA prepared by sucrose gradient centrifugation.

Detection of *unc*-54 mRNA with cDNA clones

The partially purified myosin heavy-chain mRNA was used as a template to construct recombinant plasmids containing *unc*-54-specific sequences. As expected, the large size of the *unc*-54 mRNA (~6,000 base pairs) prevented us from synthesizing full-length cDNA clones. We therefore devised a cloning strategy, based on cleavage of cDNA with restriction enzymes, which yielded short cDNA probes of the desired specificity.

cDNA was synthesized from the enriched *unc*-54 mRNA preparation using AMV reverse transcriptase selectively primed at the 3′ end with oligo(dT)[19]. The template RNA was destroyed and the cDNA made double stranded in a self-primed reaction using *Escherichia coli* DNA polymerase A[20]. Fragments of double-stranded cDNA were then generated[21] by cleavage with *Mbo*I, a nuclease which cleaves at the sequence GATC leaving a tetranucleotide extension. These short fragments (100–500 base pairs) were directly ligated into the *Bam*HI site of the plasmid pBR322[22].

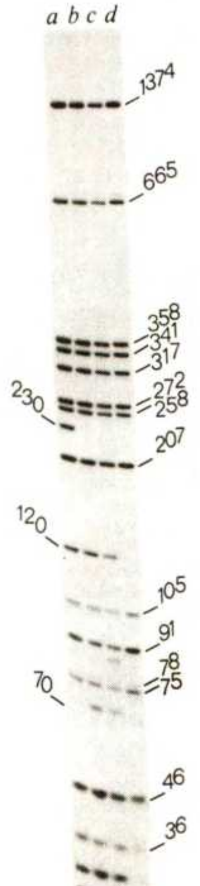
Plasmids containing *unc*-54-specific sequences were identified using an RNA blotting and hybridization assay based on the properties of wild-type and mutant *unc*-54 mRNAs (Fig. 3). The results of the cell-free translation experiments described above suggested that the *unc*-54 mRNA should be reduced in molecular weight by ~300 nucleotides in E675 animals and in low abundance in E190 animals. These differences should be detectable by hybridization of cloned *unc*-54 sequences to RNA prepared from these strains, which has been size fractionated by gel electrophoresis. Because the genetic evidence unambiguously assigns the *e*675 and *e*190 mutations to a single locus, these strain differences should only be detected when *unc*-54 sequences are used as hybridization probes.

A control experiment using ^{32}P-labelled cDNA prepared from the enriched *unc*-54 mRNA fraction demonstrates the specificity of this assay. Total RNA from wild-type N2, E675 and E190 animals and the *unc*-54 mRNA fraction were fractionated on 1.0% agarose gels and transferred to activated paper[23]. Hybridization of cDNA to the partially purified *unc*-54 mRNA fraction detects a number of RNA molecules (Fig. 3*A*). The major component of this fraction shows the expected properties of *unc*-54 mRNA. This component is detected in the total RNA fraction from wild-type animals probed with ^{32}P-labelled *unc*-54 cDNA and is reduced in size in E675; however, a similar RNA species cannot be detected in E190 RNA. This experiment provides unambiguous genetic evidence that *unc*-54 mRNA has been detected with this probe.

Figure 3*B* shows the hybridization pattern obtained when a plasmid containing *unc*-54 sequences is used as a probe in the assay (pMbo7). A single RNA species is detected in the N2 RNA fraction which shows the expected reduction in size in E675 and is absent from E190. No E190 sequences were detected with this probe, indicating either that the E190 mRNA is present below the level of detection in this experiment or that the *unc*-54-specific sequence contained in the pMbo7 probe is absent from the altered mRNA.

The structures of three plasmids containing *unc*-54 cDNA sequences are shown in Fig. 4. Plasmid DNA was digested with *Sau*3A[24] (an isoschizomer of *Mbo*I) and the restriction fragments separated by polyacrylamide gel electrophoresis. In addition to the restriction fragments of the vector, all the recombinants carry a fragment of 120 base pairs and pMbo7 and pMbo8 also carry a fragment of 70 base pairs. This suggests that these two restriction fragments are contiguous in the *unc*-54

Fig. 4 Structure of recombinant plasmids containing *unc*-54 cDNA sequences. Covalently closed circular DNA was isolated by the caesium chloride–ethidium bromide centrifugation procedure[44]. Plasmid DNA (1 μg) was digested with 2 units *Sau*3A[24] for 60 min at 67 °C in 10 mM Tris-HCl *p*H 7.4, 50 mM NaCl, 10 mM $MgCl_2$ and 10 mM DDT and was end labelled by incubation in the presence of 1 μCi [α-^{32}P]dATP, 500 μM each of nonradioactive dGTP, dCTP, dTTP and 1 unit of AMV reverse transcriptase. The labelled fragments were fractionated by electrophoresis through a thin (0.03 × 20 × 40 cm) 5% polyacrylamide gel[45] without urea in TBE buffer and detected by autoradiography of the dried gel. The sizes of the *Sau*3A restriction fragments of pBR322, based on nucleic acid sequence information[46], are indicated on the right side of the figure. The sizes of inserted *Sau*3A fragments are indicated on the left side. *a*, pMbo3; *b*, pMbo7; *c*, pMbo8; *d*, pBR322.



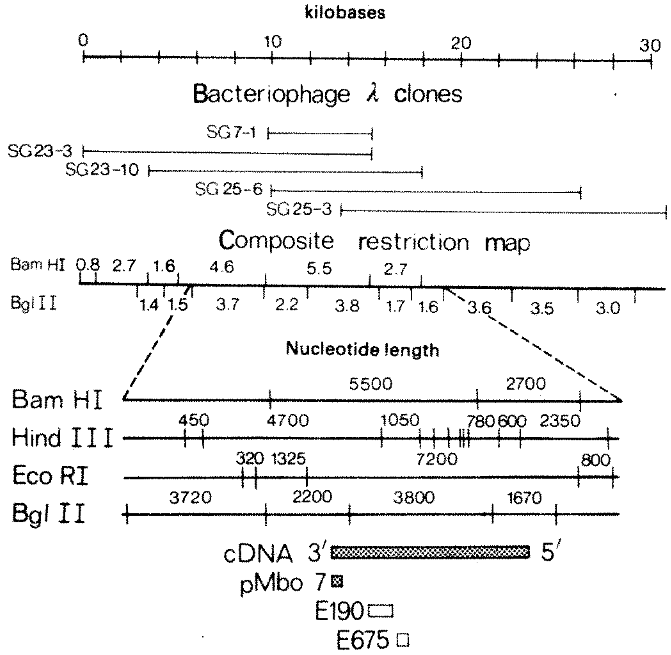

kilobases
0
10
20
30
Bacteriophage λ clones
SG7-1
SG23-3
SG23-10
SG25-6
SG25-3
Composite restriction map
Bam HI 0.8 2.7 1.6 4.6 5.5 2.7
Bgl II 1.4 1.5 3.7 2.2 3.8 1.7 1.6 3.6 3.5 3.0
Nucleotide length
Bam HI 5500 2700
Hind III 450 4700 1050 780 600 2350
Eco RI 320 1325 7200 800
Bgl II 3720 2200 3800 1670
cDNA 3′ 5′
pMbo 7
E190
E675

Fig. 5 Restriction endonuclease maps of clones containing the *unc*-54 myosin heavy chain and flanking sequences. Recombinant phages were prepared as described using bacteriophage λ1059 as a vector[30]. Representative clones are shown in the top panel. The maps are drawn without the phage sequences. SG 7-1 is a *Bam*HI clone containing the *Bam* 5.5-kilobase fragment and a second restriction fragmentation of 12-kilobases (not shown). SG 23-3 and SG 23-10 are clones obtained from *Bam*HI partial digestion of nematode DNA. They share the 5.5- and 4.6-kilobase *Bam*HI fragments but differ in their flanking sequences. SG 25-6 and SG 25-3 were obtained from clones prepared by *Sau*3A digestion of nematode DNA. They both have one end point within the *Bam* 5.5-kilobase fragment and extend leftward. A composite *Bam*HI–*Bgl*II restriction map of the 30-kilobase region defined by these overlapping clones is drawn below the phage maps. The fragment sizes are given in kilobases. A more detailed restriction map of part of this region is also drawn. The length of the fragments in nucleotides is indicated. The small unnumbered *Hin*dIII fragments have sizes of 294, 368, 357, 117 and 111 nucleotides, reading from left to right. The sites of hybridization of pMbo7 and cDNA to *unc*-54 mRNA are indicated by the cross-hatched bars. Note that the end point of the detected cDNA hybridization does not precisely define the 5′ end of the *unc*-54 mRNA sequence. The open bars indicate the positions of the *e*675 and *e*190 deletions. These positions were assigned on the basis of Southern transfer hybridization experiments of the type shown in Fig. 6.

mRNA sequence and were cloned together in pMbo7 and pMbo8 by incomplete digestion of double-stranded cDNA. The plasmid pMbo3 also carried an additional unique fragment of 230 base pairs which does not hybridize to *unc*-54 mRNA, and was probably cloned by a multiple insertion event.

*e*190 is a small deletion

The cDNA plasmids can be used to identify restriction fragments of nematode genomic DNA which contain *unc*-54 sequences. *Bam*HI-digested DNA isolated from N2, E675 and E190 animals was hybridized to nick-translated[25] pMbo7 (Fig. 6*A*) using the method of Southern[26]. A single labelled restriction fragment of 5.5 kilobases is detected in wild-type DNA. It is reduced in size in E675 by approximately 300 base pairs, consistent with the size of the internal deletion in the E675 myosin heavy chain. Surprisingly, this fragment shows a reduction in size of ~600 base pairs in E190, demonstrating that the *e*190 mutation is also a small deletion. As the sequence in the pMbo7 probe is present in the genomic DNA but is undetectable in fractions expected to contain *e*190 mRNA, this result suggests that the *e*190 mutation produces a more complicated alteration of the *unc*-54 transcription unit than might be expected to result from a simple deletion of part of the coding sequence, as in E675.

To map the *unc*-54 gene more precisely, we turned our attention to molecular cloning of genomic DNA sequences.

Molecular cloning of the *unc*-54 gene

Eukaryotic genes have been isolated from collections of randomly cloned DNA fragments[27–29]. We have recently described a novel bacteriophage λ cloning vector, λ1059, which has been used to construct a collection of recombinants covering the *C. elegans* genome[30]. High-molecular-weight N2 DNA was digested with *Bam*HI or *Sau*3A in conditions which yielded fragments of 15–20 kilobases. These were purified by agarose gel electrophoresis and ligated to the bacteriophage arms. Recombinants harbouring inserts were then selected by growth on a strain which restricted the growth of the vector.

Twenty-two recombinants which hybridized to cDNA and cDNA plasmid probes[31] were isolated and the restriction maps of each determined. These phage were found to contain a series of overlapping restriction fragments which could be ordered to generate a map extending for about 30 kilobases of the nematode genome. The composite restriction map of *Bam*HI and *Bgl*II fragments and the end points of four of these phages are shown in Fig. 5.

Orientation of the gene

The short cDNA plasmid pMbo7 hybridizes to the site within the *Bam* 5.5-kilobase fragment indicated in Fig. 5 (data not shown). The orientation of the *unc*-54 gene can be defined with respect to this site by comparing the patterns of hybridization of *unc*-54 cDNA and pMbo7. Restriction fragments to the left side of the site of pMbo7 hybridization (as the map is drawn) do not hybridize to cDNA while those to the right will hybridize efficiently (Fig. 5). Since cDNA synthesis is primed by oligo(dT) at the 3′-poly(A)-terminus of the *unc*-54 mRNA this suggests that pMbo7 contains a sequence at or near the 3′ end of the transcript. The fact that because of the very large size of the *unc*-54 mRNA very few cDNA transcripts primed with oligo(dT) were of full length introduces some ambiguity in mapping the 5′ end of the *unc*-54 coding sequence, and the map position indicated in Fig. 5 is accurate to only 300 base pairs. Preliminary evidence using mRNA as a probe in Berk–Sharp[32]

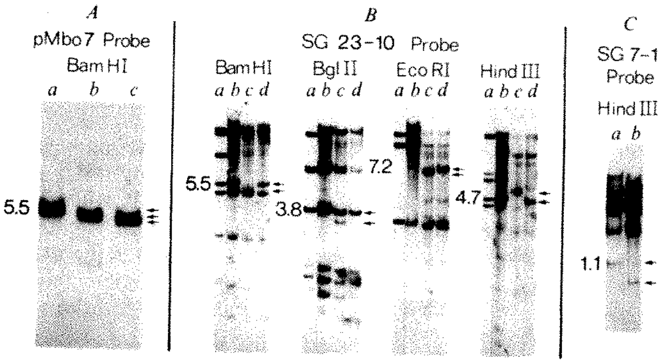


Fig. 6 Identification of *unc*-54 DNA sequences in restriction fragments of nematode genome DNA. DNA was isolated from nematodes essentially as described previously[33]. DNA (10 μg) was digested with various restriction enzymes in a 50-μl reaction containing 50 mM KCl, 10 mM Tris-HCl *p*H 7.5 and 6 mM Mg acetate and electrophoresed in agarose gel in TAE buffer (40 mM Tris-acetate *p*H 7.3, 20 mM Na-acetate, 2 mM EDTA) and transferred to nitrocellulose[26]. Bacteriophage were grown as liquid lysates on Q358 as described previously[30]. Aliquots (10 ng) of phage DNA were digested with restriction enzymes and used as molecular weight markers (*a*). Probes for Southern transfer experiments were prepared by end labelling of restriction fragments with AMV polymerase as described in Fig. 4 legend. The probe obtained from pMbo7 was a *Hin*dIII–*Sal* fragment which was purified by gel electrophoresis before hybridization to DNA immobilized on nitrocellulose. The bacteriophage probes were end-labelled *Sau*3A digests of the phage DNA. *A*, hybridization of pMbo7 to *Bam*HI-cleaved DNA from N2 (*a*), E675 (*b*) and E190 (*c*) fractionated on a 0.8% agarose gel. *B*, hybridization of SG23-10 DNA to *Bam*HI, *Bgl*II, *Eco*RI, *Hin*dIII cleaved DNA from N2 (*b*), E190 (*c*) and E1201 (*d*), fractionated on a 1% agarose gel. *C*, hybridization of SG 7-1 DNA to *Hin*dIII-cleaved DNA from N2 (*a*) and E675 (*b*) fractionated on a 1.5% agarose gel. Numbers indicate the size in kilobases of restriction fragments that are affected by the E190 and E675 deletions. These fragments are similarly identified in the restriction map in Fig. 5. Arrows indicate the positions of the altered restriction fragment in the wild-type and mutant strains.

hybridization experiments also places the 5′ end of the messenger to the indicated position within the *Bam* 2.7-kilobase fragment.

Mapping *e*675 and *e*190

We confirmed the results of this mapping experiment by more accurate mapping of the *e*675 and *e*190 deletions. As our chemical cleavage studies had shown that the terminal 300–350 amino acids of the *unc*-54 myosin heavy chain are present in E675[8], this mutation should be located approximately 1 kilobase from the end of the *unc*-54 mRNA. A similar map position should be observed in the genome, assuming that there are few intervening sequences in this region. Cell-free translation data can be used to predict the location of the *e*190 deletion. As described above, RNA isolated from E190 programmes the synthesis of a polypeptide of MW 170,000 which is immunoprecipitated by anti-myosin antiserum. A polypeptide of this size could be produced by deletion of the terminal 350 amino acid residues of the *unc*-54 myosin heavy chain. In this case, *e*190 should map near *e*675 at ~1,000 base pairs from the end of the transcribed sequence.

The Southern transfer procedure was used to map the positions of these two mutants. DNA from wild-type, E190, E675 and E1201 animals (a recessive *unc*-54 mutant) was digested with a variety of restriction enzymes and probed with radioactive SG 23-10 or SG 7-1 DNA (Fig. 6). The *e*190 deletion produces several alterations in the cleavage patterns (Fig. 6*B*, panel 1). In particular, this mutation results in the fusion of the 4.8- and 1.1-kilobase *Hin*dIII fragments found in the wild-type pattern, localizing the *e*190 deletion to the region indicated in Fig. 5. Similarly, the *e*675 deletion has been localized within the 1.1-kilobase *Hin*dIII fragment (Fig. 6*C*). The two mutations thus seem to map close to one another and may, in fact, overlap. This provides evidence for our interpretation of the E190 mRNA translation pattern and confirms that the *unc*-54 gene is oriented as shown in Fig. 5. No alteration of the restriction map of E1201 was detected in these experiments, suggesting that this recessive null mutation may be a single base pair change.

These Southern transfer experiments also confirm the restriction map derived from an analysis of the bacteriophage λ clones. No rearrangements seem to have occurred during the cloning procedure because each of the restriction fragments present in the nematode-λ recombinants are also detected in the genomic DNA.

In addition to the strongly hybridizing fragments present in the cloned probes, a number of weakly hybridizing restriction fragments were detected in these experiments. It will be of interest to determine whether any of these restriction fragments show homology to *unc*-54 sequences, and whether the other myosin heavy chains present in *C. elegans* can be detected by hybridization to *unc*-54-specific sequences in conditions of relaxed stringency.

The *unc*-54 gene

The *unc*-54 myosin heavy-chain gene of *C. elegans* provides an attractive model system for studying molecular genetics in a multicellular higher organism. Numerous mutants in the gene have been isolated, and more can be readily obtained.

Here we have described a series of complementary biochemical assays which may be used to define the molecular phenotypes of these mutations. Myosin synthesis can be studied both *in vivo* and *in vitro* by cell-free translation of the *unc*-54 mRNA and immunoprecipitation of the myosin heavy chains. Messenger RNA can be detected by hybridization to cloned probes and deletions can be mapped by Southern hybridization transfer experiments.

In developing these assays we relied on the properties of a mutant, E675, whch was known to contain an internal deletion of 100 amino acids near the carboxy terminus of the *unc*-54 myosin heavy chain[8]. Thus, *unc*-54-specific translation products, mRNA molecules and restriction fragments of the nematode genome could be identified as molecules which showed altered length in E675. This assay allowed us to identify sequences from the active gene in cloning experiments and to distinguish the products of the *unc*-54 gene from those of the other myosin heavy-chain genes present in this organism.

We have also applied these assays to E190[1] and E1201[4]. E190 was the first mutant isolated in *unc*-54[1] and was long thought to be a point mutation. However, Southern transfer experiments revealed that *e*190 is, in fact, a small deletion near the 3′ end of the *unc*-54 coding sequence. Unlike E675, E190 does not produce a stable protein product and appears deficient in *unc*-54 mRNA. However, immunoprecipitation of the myosin heavy chains synthesized *in vitro* by E190 RNA reveals the presence of a minor polypeptide specific to E190, which may be a truncated translation product of the *unc*-54 (*e*190) mRNA. The size and position of the deletion found in E190 DNA are consistent with this interpretation. In contrast to E190, E1201 shows no alteration of the genomic restriction map, suggesting that this recessive null mutation may be the result of a single base pair change. However, in other respects the phenotypes of E1201 and E190 are similar. Both mutants are deficient in *unc*-54 mRNA synthesis and both code for immunoprecipitable peptides which may be detected by cell-free translation[12].

Further analysis of the structure of the wild-type *unc*-54 gene and of these and other mutants should contribute to our knowledge of gene expression in eukaryotes as well as to the molecular biology of muscle.

We thank Dr R. P. Anderson for sharing map information about the position of the *e*190 mutation, G. P. Goelet, J. Martin and L. Barnett for helpful discussions, and R. M. Fishpool for assistance throughout this work. A.R.M. was supported by fellowships from the Muscular Dystrophy Association and the MRC of Canada. J.K. was supported by a fellowship from the Helen Whitney Foundation.

Received 22 December 1980; accepted 1 April 1981.

1. Brenner, S. *Genetics* **77,** 71–94 (1974).
2. Epstein, H. F., Waterston, R. H. & Brenner, S. *J. molec. Biol.* **90,** 291–300 (1974).
3. MacKenzie, J. M., Schachat, F. & Epstein, H. F. *Cell* **15,** 413–419 (1978).
4. MacLeod, A. R., Waterston, R. H., Fishpool, R. M. & Brenner, S. *J. molec. Biol.* **114,** 133–140 (1977).
5. Schachat, F., Harris, H. E. & Epstein, H. F. *Cell* **10,** 721–728 (1977).
6. Riddle, D. L. & Brenner, S. *Genetics* **89,** 299–314 (1978).
7. Waterston, R. H. & Brenner, S. *Nature* **275,** 715–719 (1978).
8. MacLeod, A. R., Waterston, R. H. & Brenner, S. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5336–5340 (1977).
9. Cox, R. A. *Meth. Enzym.* **12B,** 120–129 (1968).
10. Roberts, B. E. & Paterson, B. M. *Proc. natn. Acad. Sci. U.S.A.* **70,** 2330–2334 (1973).
11. Pelham, H. R. B. & Jackson, R. J. *Eur. J. Biochem.* **67,** 247–256 (1976).
12. MacLeod, A. R., Karn, J., Waterston, R. H. & Brenner, S. in *Nonsense Mutations and tRNA Suppressors* (eds Celis, J. E. & Smith, J. D.) 301–311 (Academic, New York, 1979).
13. Jacobson, G. R., Schaffer, M. H., Stark, G. R. & Vanaman, T. C. *J. biol. Chem.* **248,** 6583–6591 (1973).
14. Degani, Y. & Patchornik, A. *Biochemistry* **13,** 1–11 (1974).
15. Bailey, J. M. & Davidson, N. *Analyt. Biochem.* **70,** 75–85 (1976).
16. Aviv, H. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **69,** 1408–1412 (1972).
17. Adesnick, M., Saldith, M., Thomas, W. & Darnell, J. E. *J. molec. Biol.* **71,** 21–30 (1972).
18. Vogt, V. M. *Eur. J. Biochem.* **33,** 192–200 (1973).
19. Buell, G. N., Wickens, M. P., Payvar, F. & Schimke, R. T. *J. biol. Chem.* **253,** 2471–2482 (1978).
20. Wickens, M. P., Buell, G. N. & Schimke, R. T. *J. biol. Chem.* **253,** 2483–2495 (1978).
21. Gelinas, R. E., Myers, P. A. & Roberts, R. J. *J. molec. Biol.* **114,** 169–180.
22. Bolivar, F. *et al. Gene* **2,** 95–113 (1977).
23. Alwine, J. C., Kemp, P. J. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5350–5354 (1977).
24. Liu, A. C. P., McBride, B. C., Vovis, G. F. & Smith, M. *Nucleic Acids Res.* **6,** 1–15 (1979).
25. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. *J. molec. Biol.* **113,** 237–251 (1977).
26. Southern, E. M. *J. molec. Biol.* **8,** 503–517 (1975).
27. Wensink, P. C., Finnegan, D. J., Donelson, J. E. & Hogness, D. *Cell* **3,** 315–325 (1974).
28. Kedes, L. H., Chang, A. C. Y., Houseman, D. & Cohen, S. N. *Nature* **255,** 533–537 (1975).
29. Maniatis, T. *et al. Cell* **15,** 687–701 (1978).
30. Karn, J., Brenner, S., Barnett, L. & Cesareni, G. *Proc. natn. Acad. Sci. U.S.A.* **77,** 5172–5176 (1980).
31. Benton, W. P. & Davis, R. W. *Science* **196,** 180–182 (1977).
32. Berk, A. J. & Sharp, P. A. *Cell* **12,** 721–732 (1977).
33. Sulston, J. E. & Brenner, S. *Genetics* **77,** 95–104 (1974).
34. Kessler, S. W. *Immunology* **115,** 1617–1624 (1975).
35. Degani, Y. & Patchornik, A. *J. org. Chem.* **36,** 2727–2728 (1971).
36. Laemmli, U. K. *Nature* **227,** 680–685 (1970).
37. Bonner, W. M. & Laskey, R. A. *Eur. J. Biochem.* **46,** 83–88 (1974).
38. Laskey, R. A. & Mills, A. D. *Eur. J. Biochem.* **56,** 335–341 (1975).
39. Seeburg, P. H., Shine, J., Martial, J. A., Baxter, J. D. & Goodman, H. M. *Nature* **270,** 486–494 (1977).
40. Brownlee, G. G. & Cartwright, E. M. *J. molec. Biol.* **114,** 93–117 (1977).
41. Deeley, R. G. *et al. J. biol. Chem.* **252,** 8310–8319 (1977).
42. Cohen, S. N., Chang, A. C. Y., Boyer, H. W. & Helling, R. *Proc. natn. Acad. Sci. U.S.A.* **70,** 3240–3244 (1973).
43. Grunstein, M. & Hogness, D. S. *Proc. natn. Acad. Sci. U.S.A.* **72,** 3961–3965 (1975).
44. Clewell, D. B. *J. Bact.* **110,** 667–676 (1972).
45. Sanger, F. & Coulson, A. R. *FEBS Lett.* **87,** 107–110 (1978).
46. Sutcliffe, J. G. *Nucleic Acids Res.* **5,** 2721–2728.

LETTERS TO NATURE

Why is the Universe open?

David Lindley

Astronomy Centre, University of Sussex, Falmer,
Brighton BN1 9QH, UK

One of the most fundamental properties of the Universe is its density, or more usefully the ratio $\Omega = \rho/\rho_c$ of the density to the critical value required for closure. The feeling among cosmologists is that the Universe is open, with Ω perhaps in the range 0.1–0.5 (see ref. 1). In general, Ω is thought of as a parameter to be deduced from observation, but some reasons for it having a particular value have been suggested; there is a certain aesthetic appeal in having $\Omega = 1$ (ref. 2), and the 'anthropic principle'[3] suggests that Ω should not be too far from unity if the Universe is to evolve in such a way as to allow the development of stars, and thence ourselves. Here speculations are made concerning the transition from quantum to classical gravitational behaviour near the initial cosmic singularity, and a physical process is proposed whose action leads naturally to a Universe which is open.

In a Universe which is, on average, represented by an $\Omega = 1$ Friedmann model, an inhomogeneity in the form of a density enhancement is a region of local positive curvature. There exists a general argument[4] to say that if the density enhancement is too great, the curvature will be large enough for the region to close up on itself and be completely disconnected from the $\Omega = 1$ Universe. In classical general relativity, such disjoint regions never have been, and never will be in contact, but I suggest here that their mutual independence may be removed by quantum gravitational effects.

Quantum effects dictate that the precise definition of classical points, surfaces, and so on becomes impossible on small scales[5]. One must therefore imagine that the picture of the classical Universe as an ordered set of well-defined spacelike hypersurfaces breaks down near the Planck time, $t_p \simeq 10^{-43}$ s, as quantum fluctuations of the metric disrupt and mix up the classical surfaces[6]. The proposal made here is that the regions of density enhancement mentioned above may become physically separated, in a dynamical sense, from the rest of the Universe during the transition from quantum to classical eras: in the classical era the spacetime manifold is well-defined, and the regions are disjoint, but in the quantum era, everything is smeared together.

One can usefully envisage this process occurring in 'superspace', each point of which represents a particular spacetime geometry[5]. Classically, any cosmological model follows a trajectory of least action through superspace, but using a path integral approach to quantum gravity[7], one must consider all routes, including non-classical ones, between one geometrical configuration and the next. The present conjecture then amounts to a supposition that the Universe, as it emerges from the region of superspace where quantum gravity dominates, can reach configurations which are topologically inequivalent to Friedmann models: that is, they comprise several unconnected regions. There seems to be no very strong arguments at present for or against this hypothesis, but the consequences would be interesting.

A simple calculation will show how the density of the Universe may be determined by its initial conditions. If an inhomogeneity has, at t_p, mass m and density contrast $\delta = (\rho' - \rho)/\rho$, then the criterion that it should form a separate region[6] is

$$\delta > \delta_c = \alpha (m/m_0)^{-2/3} \qquad (1)$$

where m_0 is the horizon mass at t_p, and α is a constant of the order of one. Now suppose, following Carr[8], that the spectrum of density fluctuations is gaussian, with width $\delta_{rms} = \varepsilon (m/m_0)^{-n}$. If $n < 2/3$, inhomogeneity is greater on larger scales (in comparison with δ_c), and this case must be excluded as it does not correspond, for large m, to perturbations on a smooth background. If $n = 2/3$, inequality (1) is independent of m; this case will be mentioned later. For the moment, we assume $n > 2/3$, in which case fluctuations with $m = m_0$ are most important, and those with $m > m_0$ can be neglected.

The fraction of fluctuations with $\delta > \delta_c$ is roughly[8] $P = \exp(-x^2)/x$, where $x = \alpha/\sqrt{2}\varepsilon$. P can be taken to be the fraction of matter which is 'lost' in the quantum to classical transition. (If one supposed the fluctuations not to be spherically symmetric, and that P were the probability of the curvature closing on itself in one space direction, then it could be argued that P^3 would be the fraction of matter lost. This in fact makes little difference).

If Ω_0 is the cosmic density ratio today, and Ω_p its value at t_p, then it is easy to show that

$$1 - \Omega_0 = (1 - \Omega_p)(t_{eq}/t_p)(t_0/t_{eq})^{2/3} \qquad (2)$$

where the Universe is radiation dominated from $t \simeq 10^{-43}$ s to $t_{eq} \simeq 10^{12}$ s, and matter dominated from t_{eq} to the present, $t_0 \simeq 10^{18}$ s. If the Universe is noticeably open now ($\Omega_0 < 1$), then $\Omega_p \simeq 1 - 10^{-60}$. We now have a possible explanation for the tiny, but non-zero, departure from the $\Omega = 1$ 'ideal': a small fraction of the cosmic density was lost at the beginning. Taking P (or P^3) to be 10^{-60} requires $\alpha/\varepsilon = 15$ (or 10), which seems plausible; $\varepsilon = 1$ represents a chaotic initial state, with lower values signifying a more uniform Universe, and it might be expected that the simple analysis leading to equation (1) underestimates the difficulty with which separation of regions occurs, so that α might be rather larger than one.

In the particular case $n = 2/3$, the probability of regions becoming separate is independent of their size; thus one might imagine a whole 'family' of Universes emerging from the quantum regime. However, there is a problem in that the observed large-scale structure of the Universe suggests that the magnitude of $n = 2/3$ fluctuations is perhaps $\varepsilon \simeq 10^{-4}$ (ref. 9), so that the probability of there being any disjoint regions would be very small ($P \simeq \exp(-10^8)$).

It would be a mistake to take this calculation seriously in any quantitative way, and although the numbers fit together quite well, we cannot claim to have predicted the density of the Universe because the path of cosmological evolution is extremely sensitive to the value of Ω_p. In fact, if α/ε has a value of 1 rather than 10, a large fraction of the density will be lost at t_p, and the Universe will expand and recontract far too quickly to allow stars and galaxies to develop. Low values of α/ε clearly do not lead to Universes like our own, and although a reason has been suggested why $\Omega < 1$, there still seems to be some sort of anthropic coincidence in the fact that Ω_0 is only slightly less than 1 after 10^{10} yr. Another puzzle lies in finding a sounder physical basis for the assumption that $\Omega = 1$ is a 'preferred' state; the idea that such a Universe has an exact balance between kinetic and gravitational energy[2] is interesting, but, at present, no more than suggestive.

Recently, Guth[10] has looked at a possible explanation for some of the gross features of the Universe, but the mechanism he uses (phase transitions in Grand Unified Theories) is completely different from the one discussed here; therefore no direct comparison of the two approaches seems possible. Nevertheless, it is a new and intriguing idea that the density of the Universe may be, as the cosmic baryon number has recently been shown to be, a predictable quantity.

I thank Dr B. J. Carr and Professor R. J. Tayler for sympathetic criticism, and the SRC for a research studentship. Dr C. J. Hogan made several useful comments.

0028-0836/81/220391—02$01.00

Received 2 February; accepted 14 April 1981.

1. Gott, J. R. III, Gunn, J. E., Schramm, D. N. & Tinsley, B. M. *Astrophys. J.* **194,** 543–554 (1974).
2. Dicke, R. H. & Peebles, P. J. E. in *General Relativity: an Einstein Centenary Survey* (eds Hawking S. W. & Israel, W.) 504–517 (Cambridge University Press, 1979).
3. Carter, B. in *Confrontation of Cosmological Theories with Observational Data* (ed. Longair, M. S.) 291–298 (Reidel, Dordrecht, 1974).
4. Carr, B. J. & Hawking, S. W. *Mon. Not. R. astr. Soc.* **168,** 399–415 (1974).
5. Misner, C. W., Thorne, K. S. & Wheeler, J. A. *Gravitation* (Freeman, San Francisco, 1973).
6. Harrison E. R. *Phys. Rev.* **D1,** 2726–2730 (1970).
7. Hawking, S. W. in *General Relativity: an Einstein Centenary Survey* (eds Hawking, S. W. & Israel, W.) 746–789 (Cambridge University Press, 1979).
8. Carr, B. J. *Astrophys. J.* **201,** 1–19 (1975).
9. Gott, J. R. III *A. Rev. Astr. Astrophys.* **15,** 235–266 (1977).
10. Guth, A. H. *Phys. Rev.* **D23,** 347–356 (1981).

IR variability of SS433

R. M. Catchpole, I. S. Glass*, B. S. Carter & G. Roberts

South African Astronomical Observatory, PO Box 9,
Observatory 7935, Cape, South Africa

SS433 shows X-ray, optical and radio variability and evidence of material moving at one-quarter the speed of light[1]. We report here the presence of the 165.5-day spectroscopic period in our extensive *JHKL* (1.2–3.4 μm) photometry. This indicates that the jets, rather than a highly reddened O star, are a major source of the IR radiation. We also show that the $(J-H)$ colour index becomes redder as SS433 brightens but find no evidence for the 11.8-day period reported by Giles *et al.*[2].

Seventy-two *JHK* observations of SS433 have been made at the South African Astronomical Observatory with the 1.88- and 0.75-m telescopes. The observations span three observing seasons and include a run in 1980, when the star was observed on 33 out of 55 consecutive nights. The data and standard errors are listed in Table 1. The errors contain an 'internal' standard error of observation and an 'external' error which is fixed at $\Delta J = \pm 0.06$, $\Delta H = \pm 0.05$, $\Delta K = \pm 0.04$ and $\Delta L = \pm 0.04$ for the 0.75-m telescope and at ±0.03 for all filters on the 1.88-m telescope. We list *L* magnitudes obtained only with the 1.88-m telescope although other less accurate ones, made with the 0.75-m telescope, are available. The observations have all been reduced to the system of Glass[3] standards. On six occasions in 1979 Giles *et al.* and ourselves observed SS433 within 6 h of each other. For these nights our observations are systematically brighter by 0.13 in *J*, and 0.06 mag in *H* and *K*. The simultaneous observations cover a range of *K* magnitudes from 8.3 to 7.2 and so independently confirm the reality of the rapid change in brightness which we have observed. We have not applied any systematic corrections to the data from Giles *et al.* because we cannot assume that such corrections are universally applicable o their rather heterogeneous observations.

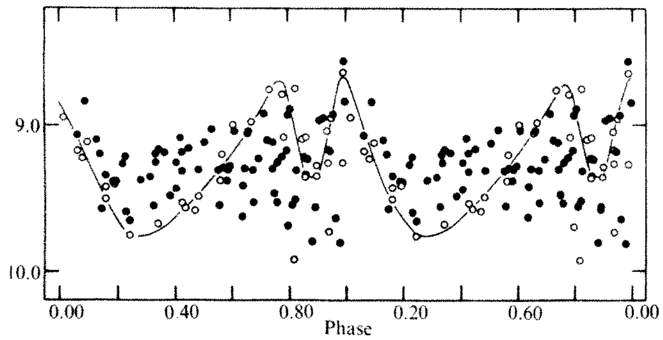


Fig. 1 The *J* magnitude as a function of phase for an 11.8-day period. ●, Data in Table 1. ○, Data given by Giles *et al.* The curve is their proposed light curve. Phase 0.0 is on JD 2443718.9.

A sequential plot of the data shows that the brightness in all colours fluctuates by tenths of a magnitude on a time scale of days while occasional brightening and fadings of 0.7 mag occur on the time scale between 4 and 5 days. Clearly any underlying periodicities will probably be obscured by this 'flickering'. Period-finding programmes must therefore be used with great care as the occurrence of a few strong brightenings, coupled with the very patchy distribution of data, can cause spurious periodicities.

Table 1 *JHKL* observations of SS433

JD	*J*	± Δ*J*	*H*	± Δ*H*	*K*	± Δ*K*	*L**	± Δ*L**
3729.40	9.56	0.07	8.94	0.05	8.19	0.04		
3730.43	9.81	0.07	9.19	0.06	8.44	0.07		
3732.43	9.58	0.06	9.03	0.07	8.28	0.05		
3733.43	9.60	0.06	9.01	0.05	8.35	0.05		
3999.64*	9.17	0.03	8.50	0.03	7.84	0.03		
4001.55*	8.93	0.03	8.27	0.03	7.64	0.03		
4003.63	9.10	0.06	8.47	0.05	7.84	0.04		
4070.49	9.69	0.07	9.12	0.05	8.33	0.04		
4071.48	9.80	0.08	9.10	0.06	8.29	0.05		
4072.44	9.64	0.08	8.91	0.05	8.26	0.04		
4079.48*	9.55	0.03	8.90	0.03	8.20	0.03		
4080.41	9.63	0.11	8.87	0.05	8.27	0.06		
4082.46	9.55	0.08	8.86	0.07	8.09	0.09		
4083.40*	9.56	0.03	8.91	0.03	8.21	0.03		
4083.41	9.57	0.08	9.06	0.06	8.22	0.06		
4094.39	9.51	0.08	8.90	0.05	8.31	0.16		
4096.39	8.56	0.06	7.78	0.05	7.08	0.04		
4099.39	9.66	0.06	9.08	0.06	8.27	0.05		
4100.37	9.56	0.06	8.95	0.05	8.23	0.04		
4101.31	9.44	0.08	8.75	0.09	8.19	0.04		
4103.40	9.30	0.06	8.60	0.05	7.96	0.04		
4107.37	8.95	0.08	8.18	0.05	7.56	0.04		
4109.37	8.84	0.07	8.12	0.05	7.45	0.04		
4113.35	9.32	0.07	8.70	0.06	8.12	0.12		
4140.30*	8.92	0.03	8.17	0.03	7.54	0.03		
4141.29*	8.93	0.03	8.22	0.03	7.61	0.03		
4141.38*	8.89	0.03	8.22	0.03	7.59	0.03		
4171.25*	9.20	0.03	8.55	0.03	7.94	0.03		
4172.25*	9.09	0.03	8.41	0.03	7.80	0.03		
4173.26*	9.12	0.03	8.42	0.03	7.83	0.03		
4175.25*	9.31	0.03	8.60	0.03	7.94	0.03		
4176.25*	9.53	0.03	8.85	0.03	8.18	0.03		
4317.65*	9.11	0.03	8.44	0.03	7.85	0.03		
4346.61	9.39	0.07	8.79	0.05	8.11	0.05		
4347.64	9.38	0.06	8.64	0.06	8.02	0.05		
4348.62	9.19	0.06	8.48	0.06	7.83	0.04		
4349.62	9.16	0.06	8.52	0.05	7.89	0.04		
4360.65	9.49	0.07	8.88	0.06	8.19	0.05		
4362.64	9.31	0.07	8.64	0.06	8.02	0.05		
4363.64	9.42	0.07	8.79	0.06	8.14	0.05		
4429.51	9.27	0.06	8.57	0.06	7.84	0.04		
4434.52	9.30	0.06	8.58	0.05	7.96	0.04		
4435.49	9.10	0.06	8.36	0.05	7.70	0.04		
4439.44	9.07	0.06	8.40	0.05	7.73	0.04		
4440.38	9.20	0.06	8.58	0.05	7.88	0.04		
4441.42	9.22	0.06	8.59	0.05	7.84	0.04		
4442.43	9.36	0.06	8.69	0.05	7.97	0.04		
4444.42	9.31	0.06	8.66	0.05	7.96	0.04		
4445.45	9.29	0.06	8.56	0.05	7.89	0.04		
4446.48	9.04	0.06	8.32	0.05	7.59	0.04		
4447.45	9.30	0.06	8.61	0.05	7.93	0.04		
4448.43	9.31	0.07	8.64	0.05	7.92	0.04		
4449.43	8.96	0.06	8.23	0.05	7.52	0.05		
4450.45	8.84	0.06	8.05	0.05	7.27	0.05		
4452.40	9.35	0.06	8.65	0.05	7.96	0.04		
4454.47	9.26	0.06	8.59	0.06	7.96	0.05		
4455.30	9.26	0.06	8.60	0.05	7.96	0.04		
4457.38	9.38	0.07	8.80	0.06	8.06	0.05		
4458.47	9.53	0.07	8.79	0.06	8.09	0.04		
4459.33	9.47	0.07	8.72	0.06	8.14	0.05		
4466.37*	9.17	0.03	8.47	0.03	7.88	0.03		
4467.33*	9.19	0.03	8.45	0.03	7.85	0.03	7.18	0.03
4469.31*	9.28	0.03	8.56	0.03	7.92	0.03	7.14	0.03
4471.28	9.25	0.07	8.51	0.05	7.92	0.05		
4472.39	9.22	0.07	8.49	0.05	7.88	0.05		
4.473.34	9.16	0.07	8.50	0.05	7.84	0.05		
4476.35	9.38	0.06	8.72	0.07	8.07	0.04		
4480.36*	9.03	0.03	8.36	0.03	7.80	0.03		
4481.29*	9.04	0.03	8.36	0.03	7.77	0.03	7.10	0.03
4482.29*	9.23	0.03	8.56	0.03	7.95	0.03	7.31	0.03
4483.23*	9.22	0.03	8.57	0.03	7.96	0.03		
4484.29*	9.23	0.03	8.52	0.03	7.93	0.03	7.32	0.03

* 1.88-m telescope observations.

---

* Present address: European Southern Observatory, 2 Karl Schwarzschild Street, D-8046, Garching, FRG.

0028-0836/81/220392—03$01.00

On Fig. 1, our data are folded onto the Giles *et al.*-suggested 11.8-day period, where the curve is their adopted lightcurve. For clarity the points are all displayed twice, covering two cycles. The more extensive data now available do not confirm their period. It is premature to use the IR data to define periodicities in the SS433 system, but it is important to see whether the well defined 165.5-day spectroscopic period, or other suggested spectroscopic periods are present in our data.

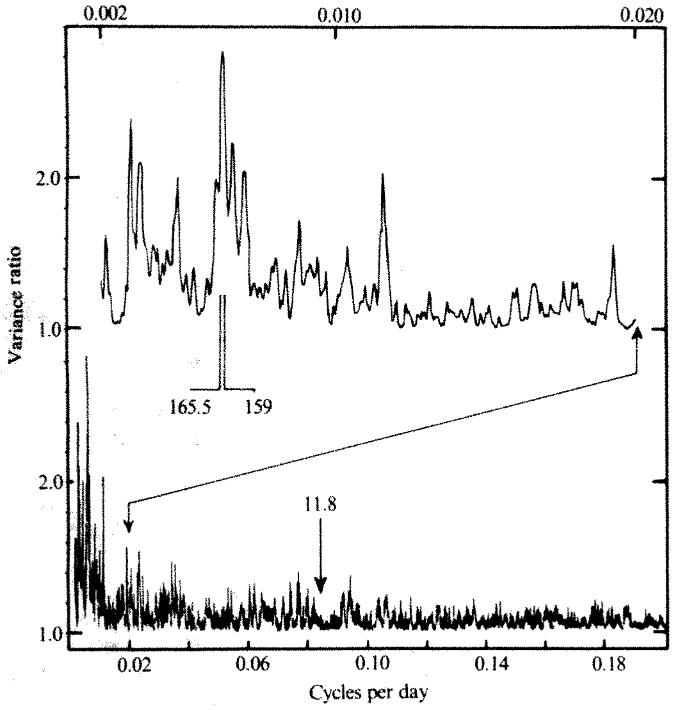


Fig. 2 The variance ratio using 5 bins for the SAAO data fainter than 7.6 K magnitude, is shown as a function of frequency, for periods in the range 5–500 days. The method is described in the text. The upper section shows the interval 50–500 days on an expanded frequency scale. Note the peak at 159 days and the absence of any peak at 11.8 days.

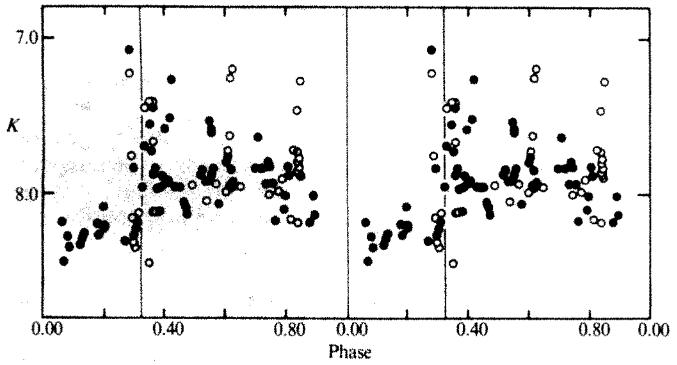


Fig. 3 The K magnitude as a function of phase for a 165.5-day period. Symbols are as in Fig. 1. The three vertical lines indicate the phases at which the radial velocities of the two high velocity emission line systems are equal. Phase 0.0 is on JD 2443718.9.

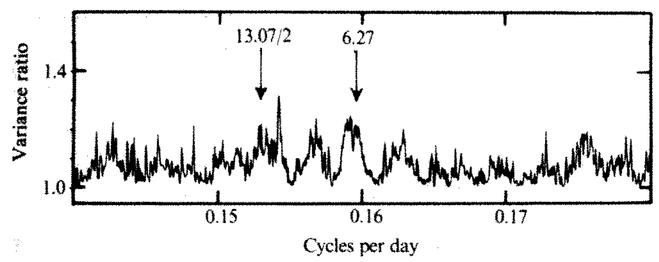


Fig. 4 The variance ratio is shown as a function of frequency for the faint SAAO data after removal of the 165.5-day period. The period interval is 5.5–7.0 days. Our weak peak at 6.27 days is marked as well as the position of Kemp *et al.* 13.07/2 period.

To avoid the effects of flaring we confine our search to data fainter than 7.6 in K. We use a method of periodic finding which, for each frequency, divides the data into phase bins and forms the ratio of the variance of the unbinned data to the weighted mean variance of the binned data. The weight is assigned according to the number of points in each bin. The method is especially suited to finding non-sinusoidal periodicities. Using this method with five bins and searching between 2 and 500 days as illustrated in Fig. 2 indicates that the most significant periodicity occurs at 159.0 days, with a variance ratio of 2.84. Snedecor's F test indicates that such a variance ratio has only a 1% probability of occurring by chance. Although it is difficult to assign quantitative errors to periods, the uncertainty can be judged from the breadth of the variance ratio peak which is ±8 days at a variance ratio height of 2.0. This suggests that our periodicity is identical to the very well defined spectroscopic period of 165.5 days given by Ciatti *et al.*[4] for the high-velocity emission line system. Figure 3 shows the data folded onto a 165.5-day period. Note that although there is no evidence for the 165.5-day period in the Giles *et al.* data alone, their data are not incompatible with this period. The three vertical lines in Fig. 3 indicate the phases at which the two high-velocity emission line velocity curves cross. In the jet model this corresponds to the phase when both jets are perpendicular to the line of sight. A striking feature of this plot is that the object is faint between phases 0.0 and 0.2. Note that flickering does not occur between these phases. The 165.5-day period has a range of 0.47 mag and shows excellent phase and amplitude agreement with the V magnitude periodicity found by Kemp *et al.*[5].

Removing the 165.5-day periodicity from our data gives evidence for a 6.27-day period with a range of ~0.2 mag, an uncertainty of ±0.03 days and a variance ratio of only 1.22 which is illustrated in Fig. 4. This period is very poorly defined but is probably to be identified with the 13.07-day double minimum periodicity found by Kemp *et al.* This in turn is identical with the 13.1-day periodic variation in the velocity of the stationary line spectrum found by Crampton *et al.*[6] and thought to represent the motion of the central binary system in SS433. Confirmation of the reality of the 6-day period in our data and better phase coverage of the 165-day period must await the results of a further season's observing.

Figure 5 shows the $(J-H)$ colour index plotted as a function of the K magnitude and fitted by a least-squares line. This shows a definite trend towards redder colour, as the star becomes brighter contrary to the conclusions of Giles *et al.* However, this trend

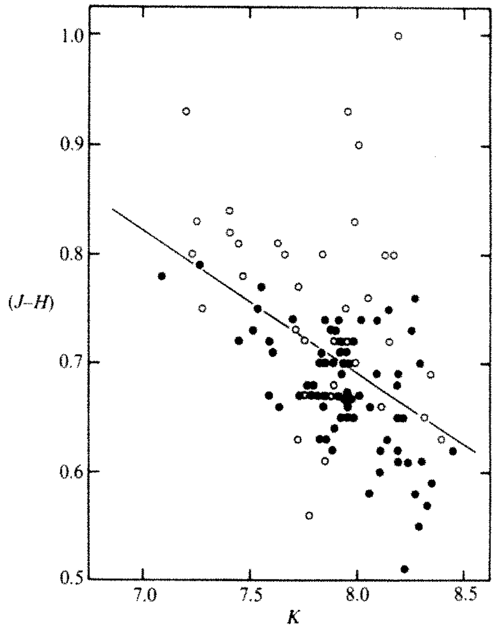


Fig. 5 The line, which is a least squares fit to all the data, illustrates the tendency of the $(J-H)$ colour to become redder as the star becomes brighter. Symbols are as in Fig. 1. The line has slope -0.129 ± 0.025.

is not seen in (*H–K*). Wynn-Williams and Becklin[7] find that in the 1 – 10-μm region SS433 becomes redder as it gets brighter, the trend increasing towards longer wavelengths. The (*J–H*) variation we observe may be related to the variation in the emission intensity and wavelength of H Paschen β, rather than to the changing optical depth in the free–free emission plasma, proposed by Wynn-Williams and Becklin to explain their data.

Murdin *et al.*[8] suggest that the continuum radiation from a highly reddened O star contributes an important part of the *JHK* flux. If this were the case we might expect the 12-day period, which is considered the binary period of the system, to be the most conspicuous. It is also not clear how their model could explain the variation of (*J–H*) with magnitude. On the other hand, McAlary and McLaren[9] consider that the *JHKL* region is dominated by variable, line and optically thin free–free emission. The 165.5-day period presumably indicates that a major part of the *JHK* flux must originate from some part of the system involved with the jets or disk. The fact that the periodicities and trends of colour with magnitude are obscured by much 'noise' suggests that we are observing radiation generated by several mechanisms.

We thank Drs M. W. Feast, J. G. Kemp and P. Murdin for valuable comments and Dr L. Balona for computer programs.

Received 2 December 1980; accepted 6 April 1981.

1. Petterson, J. A. *Rev. at 23rd COSPAR Conf.* Budapest (1980).
2. Giles, A. B. *et al. Nature* **286**, 689–691 (1980).
3. Glass, I. S. *Mon. Not. astr. Soc. S. Afr.* **33**, 53–58 (1974).
4. Ciatti, F., Mammano, A. & Vittone, A. Preprint (Series N Osservatorio Astronomico di Padova 1980).
5. Kemp, J. C., Barbour, M. S., Kemp, G. N. & Hagood, D. M. *Vistas Astr.* (submitted).
6. Crampton, D., Cowley, A. P. & Hutchings, J. B. *Astrophys. J. Lett.* **235**, L131–135 (1980).
7. Wynn-Williams, C. G. & Becklin, E. E. *Nature* **282**, 810–811 (1979).
8. Murdin, P., Clark, D. H. & Martin, P. G. *Mon. Not. R. astr. Soc.* **193**, 135–151 (1980).
9. McAlary, W. C. & McLaren, R. A. *Astrophys. J.* **240**, 853–858 (1980).

Thermal continuum emission from magnetic rotators in quasars

Wilfred H. Sorrell

625 North Frances Street, No. 9, Madison, Wisconsin 53703, USA

Although accretion of gas onto supermassive black holes is a promising explanation of the energy source for quasars[1], magnetic rotator models[2] not only satisfy the large energy requirements, but they also can explain the X-ray emission from both quasars and Seyfert galaxies[3]. It is important to compare theoretical predictions of these two models with observations so that we may determine which model is correct. Shields[4] has already done this for the black-hole accretion model; he finds that the thermal continuum emission from an accretion disk orbiting a $10^8\ M_\odot$ black hole can explain the flat blue continuum of 3C273 and other quasars. I argue here that this flat continuum also finds a natural explanation on the magnetic-rotator model.

Supermassive rotators can be represented by $n = 3$ polytropes in which radiation pressure dominates the gas pressure throughout the body. Their properties have been discussed elsewhere[5–7]. A polytrope of index n is a self-gravitating configuration in which the pressure distribution P obeys the law $P = K\rho^\gamma$, where ρ is the mass density distribution, K is a constant, and $\gamma = 1 + 1/n$ with $n > 0$. We are concerned with the (almost spherical) high-entropy configurations in equilibrium hydrogen-burning phases. In quasihydrostatic equilibrium, the central density is

$$\rho_c = 3A_n M/4\pi R_p^3 \qquad (1)$$

and the central pressure is

$$P_c = W_n GM^2/R_p^4 \qquad (2)$$

where M is the total mass and R_p is the polar radius (for polytropes of index $n = 3$, Chandrasekhar[8] gives $A_n = 54.1825$ and $W_n = 11.05066$). As radiation pressure dominates the gas pressure, we may write the central pressure

$$P_{rad} = \tfrac{1}{3}aT_c^4 = W_n GM^2/R_p^4 \qquad (3)$$

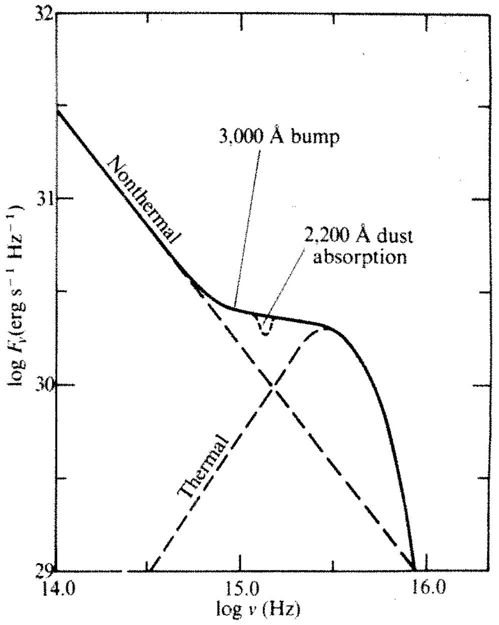


Fig. 1 The model continuum for $M_7 = 6$ and $Z_0 = 1$. The solid curve indicates the total continuum emission and a dust absorption feature near $\lambda = 2{,}200$ Å is drawn. The nonthermal power $L_{em} = 5.8 \times 10^{48}$ erg s^{-1} and the magnetic field $B \sim 10^3$ G.

Hence the radius

$$R_p = \left(\frac{3W_nG}{a}\right)^{1/4} \frac{M^{1/2}}{T_c} \qquad (4)$$

For supermassive rotators, the central temperature is fixed by the condition for thermal energy balance in which the rate of thermal energy loss equals the rate of nuclear energy generation on the carbon–nitrogen cycle. As the bulk of the nuclear reactions occur within a small temperature range, we may follow Schwarzschild[9] and write the energy generation rate per unit mass

$$\varepsilon_{cno} = \varepsilon_1 \rho_c X X_{cn} T_c^{13} \qquad (5)$$

where $X = 0.7$ is the hydrogen abundance, $X_{cn} = Z/3$ is the abundance of carbon and nitrogen nuclei, Z is the total metal abundance, and $\varepsilon_1 = 3 \times 10^{-91}$ in cgs units.

It has been shown[5,6] that supermassive bodies in general would radiate their thermal energy at a rate approximating that of the Eddington luminosity $L_{Edd} = 4\pi cGM/\kappa$, where κ is the electron scattering opacity. Thus, the condition for thermal energy balance is approximately

$$\varepsilon_{cno} M = \varepsilon_1 X \frac{Z}{3} \rho_c M T_c^{13} = 4\pi cGM/\kappa \qquad (6)$$

Using equations (1), (4), and (6) gives the central temperature

$$T_c = 3 \times 10^7\, Z_0^{-1/16} \left(\frac{M}{M_\odot}\right)^{1/32} \quad \mathrm{K} \qquad (7)$$

where the reference abundance $Z_0 = Z/Z_{cosmic}$ and the cosmic abundance $Z_{cosmic} = 0.03$. Thus, the equilibrium polar radius is

$$R_p = 1.8 \times 10^{11}\, Z_0^{1/16} \left(\frac{M}{M_\odot}\right)^{15/32} \quad \mathrm{cm} \qquad (8)$$

from equations (4) and (7).

The equatorial radius R_* is slightly larger than R_p as a result of rotary and magnetic forces at the equator. But as (high entropy) magnetic rotators are almost spherical objects[2], we shall take $R_* = 1.5\ R_p$ without any serious loss of generality. Thus

$$R_* = 5.25 \times 10^{14}\, Z_0^{1/16} M_7^{15/32} \quad \mathrm{cm} \qquad (9)$$

where M_7 is M measured in units of $10^7\ M_\odot$.

0028-0836/81/220394—03$01.00

A magnetic rotator (in its hydrogen-burning phase) would radiate a steady thermal luminosity at the Eddington rate

$$L_{\text{Edd}} = 4\pi cGM\kappa^{-1} = 1.25 \times 10^{45}\, M_7 \quad \text{erg s}^{-1} \qquad (10)$$

for about 10^6 yr. The effective temperature of the rotator (that is, that temperature the rotator would have if it radiated like a black body) is

$$\theta_* = \left(\frac{L_{\text{Edd}}}{4\pi\sigma R_*^2}\right)^{1/4} = 5 \times 10^4\, Z_0^{-1/32}\, M_7^{1/64} \quad \text{K} \qquad (11)$$

where σ is Stefan's constant. Hence the thermal continuum emission peaks in the UV parts of the spectrum near the frequency

$$\nu_0 = 2.8\, k\theta_*/h = 2.9 \times 10^{15}\, Z_0^{-1/32}\, M_7^{1/64} \quad \text{Hz} \qquad (12)$$

Note that the mass–radius relation $R_* \propto M^{15/32}$ for supermassive rotators leaves the characteristic frequency $\nu_0 \propto L_{\text{Edd}}^{1/4} R_*^{-1/2} \propto M^{1/64}$ only weakly dependent on gross properties of the rotator (such as mass). This implies that ν_0 is almost the same $\sim 10^{15}$ Hz for a wide range of masses, a property that need not be the case for arbitrary gaseous configurations in quasistatic equilibrium.

If the rotator carries a magnetic-dipole field, then emission of magnetic-dipole radiation at the speed of light cylinder would accelerate particles to high energies there[2,10,11]. This process occurs such that the rate of energy gain balances the synchrotron and Compton energy-loss rates[11]. For a rotator spinning with angular velocity

$$\Omega_* \cong \left(\frac{GM}{R_*^3}\right)^{1/2} = 3 \times 10^{-6}\, Z_0^{-3/32}\, M_7^{-13/64} \quad \text{s}^{-1} \qquad (13)$$

the speed of light cylinder has radius

$$R_1 = \frac{c}{\Omega_*} = 9.8 \times 10^{15}\, Z_0^{3/64}\, M_7^{13/64} \quad \text{cm} \qquad (14)$$

so emission variations would occur on time scales $\sim\Omega_*^{-1} \sim 4$ days. The energy flow across area $4\pi R_1^2$ carries an electromagnetic flux $\sim cB^2/4\pi$, where B is the magnetic field at radius R_1. Hence the electromagnetic power

$$L_{\text{em}} \cong B^2 c R_1^2 = 2.8 \times 10^{48} \left(\frac{B}{10^3\ \text{G}}\right)^2 Z_0^{3/32}\, M_7^{13/32} \quad \text{erg s}^{-1} \qquad (15)$$

Neugebauer *et al.*[12] have reported the spectral energy distribution from 3,000 Å to 10 μm for about 40 quasars. In general the continua can be adequately represented by a power law $f_\nu \propto \nu^{-1}$ in the IR, but there are significant deviations from power-law behaviour in the optical and UV. I propose here that the enhanced blue continua of quasars is black-body emission from a central magnetic rotator; and that the characteristic frequency associated with this thermal emission is that given by equation (12).

It has been shown[11] that synchrotron emission from radius R_1 shows a power-law spectrum

$$F_s = 6 \times 10^{-21}\, NB \left(\frac{\nu_s}{\nu}\right)^{\alpha} \quad \text{erg s}^{-1}\ \text{Hz}^{-1} \qquad (16)$$

for $\alpha > 0$ and $\nu_s = 2.83 \times 10^6\ \text{B}\,(E/mc^2)^2$ Hz. N is the total number of cosmic-ray electrons with energy $E > mc^2$.

From the magnetic rotator I predict a black-body spectrum

$$F_\nu^* = 3.9 \times 10^{30}\, Z_0^{1/32}\, M_7^{63/64} \frac{(\nu/\nu_0)^3}{\exp(h\nu/k\theta_*) - 1} \quad \text{erg s}^{-1}\ \text{Hz}^{-1} \qquad (17)$$

Figure 1 shows the model composite spectrum (non-thermal and thermal) for a rotator with $M = 6 \times 10^7\, M_\odot$ and $Z_0 = 1$. The non-thermal spectrum $F_s(\nu) = 5 \times 10^{29}\, (\nu_0/\nu)^{6/5}$ erg s^{-1} Hz^{-1}, which corresponds to model parameters $\alpha = 6/5$, $\nu_s = \nu_0 = 3 \times 10^{15}$ Hz, $B \cong 10^3$ G, $E = 10^3\, mc^2 = 5 \times 10^8$ eV, and $N \cong 8.3 \times 10^{46}$. The illustration clearly shows that the synchrotron spectrum dominates the IR, but the black-body emission dominates the UV near the Lyman continuum. On the magnetic-rotator model it is the combination of nonthermal emission from the light cylinder and thermal emission from the rotator that produces a flattening of the continuum in the blue. The model also predicts that the spectrum of ionizing continua in quasars and Seyfert galaxies should closely follow a black-body law rather than the power-law behaviour ($F_\nu \propto \nu^{-1}$) assumed in most photoionization models.

Richstone and Schmidt[13] reported a wave-like feature in the UV continuum of 85 quasars. There are spectral minima near 4,000 Å and 2,000 Å, and a continuum maximum at 3,000 Å. To interpret this behaviour on the present model, I propose that the minimum near 4,400 Å marks the onset of the thermal continuum excess itself (see Fig. 1) whereas the minimum near 2,000 Å is produced by absorbing dust particles surrounding the central rotator. Hence the intermediate emission excess should appear as a 3,000 Å bump in the spectrum. The characteristic frequency $\nu_0 \propto M^{1/64}$ associated with this bump is almost the same $\sim 10^{15}$ Hz in quasars showing quite different redshifts. Such a uniform spectral property is unexplained on black-hole accretion models.

Figure 2 shows the model composite continuum in comparison with the near IR, optical, and UV flux data[14] for the quasar 3C273. The magnetic rotator has mass $M = 7 \times 10^8\, M_\odot$ and $Z_0 = 1$. The synchrotron spectrum is the same as that for Fig. 1, so the model parameters for 3C273 are $E = 10^3\, mc^2$, $B \cong 10^3$ G, and $N \cong 8 \times 10^{46}$. The power-law spectrum dominates the IR, but there is a large flux excess in the optical and UV. Although a continuum depression is observed near 2,000 Å, the model predicts that the continuum should rise into the far UV and then turn over at frequencies between $10^{15.5}$ and 10^{16} Hz. The actual depression near 2,000 Å might be caused by absorbing dust particles or by absorption lines produced by gaseous material.

Several authors[14,15] agree that Balmer continuum emission from recombining ionized hydrogen cannot explain the bump. The same conclusion applies to two-photon continuum emission from H I gases. It has been suggested[16] that the bump is caused by emission lines produced by dust grains surrounding the broad emission-line regions of quasars. Notice, however, that the 3,000 Å bump in the spectrum of 3C273 has a luminosity $\sim 9 \times 10^{45}$ erg s^{-1}. Such a high luminosity might require an anomalously large amount of dust that would be incompatible with *UBV* colours of quasars in general.

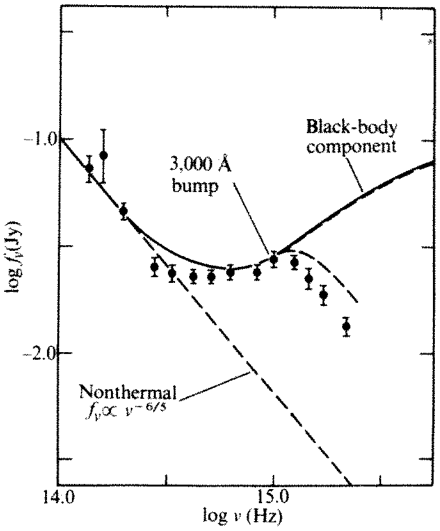


Fig. 2 The model continuum for 3C273. The solid curve indicates the model continuum flux for Hubble constant $H_0 = 100$ km s^{-1} Mpc^{-1} and distance 512 Mpc at redshift 0.158 in an empty ($q_0 = 0$) Friedmann Universe. The dashed curve near 2,000 Å indicates the continuum depression that could be caused by absorbing gases or dust. The IR, optical, and UV flux data (without reddening corrections) were taken from ref. 14. The nonthermal power $L_{\text{em}} \cong 10^{49}$ erg s^{-1} whereas the magnetic field $B \cong 10^3$ G. $M = 7 \times 10^8\, M_\odot$; $\nu_0 = 3 \times 10^{15}$ Hz.

Grandi and Phillips[15] suggested that the bump is caused by emission-line blends (that is, Fe II multiplets). Although blends of emission lines are likely to make some contribution to the 3,000 Å bump, they alone cannot explain the whole behaviour of quasar spectra in the UV. Recent IUE data[17] show that the flat spectrum continues until $10^{15.4}$ Hz, and thereafter the spectrum shows a marked steepening at higher frequencies. This spectral behaviour is found in many Seyfert galaxies and high-redshift quasars. The present model shows it in Fig. 1.

I thank Dr B. Savage for helpful advice.

Received 11 September 1980; accepted 1 April 1981.

1. Rees, M. *Ann. N.Y. Acad. Sci.* **302,** 613 (1977).
2. Ginzburg, V. L. & Ozernoy, L. M. *Astrophys. Space Sci.* **50,** 23 (1977).
3. Pacini, F. & Salvati, M. *Astrophys. J. Lett.* **225,** L99 (1978).
4. Shields, G. *Nature* **272,** 706 (1978).
5. Fowler, W. A. *Astrophys. J.* **144,** 180 (1966).
6. Fricke, K. J. *Astrophys. J.* **189,** 535 (1974).
7. Ozernoy, L. M. & Usov, V. V. *Astrophys. Space Sci.* **13,** 3 (1971).
8. Chandrasekhar, S. *Stellar Structure* (University of Chicago Press, 1938).
9. Schwarzschild, M. *Structure and Evolution of the Stars* (Dover, New York, 1958).
10. Ozernoy, L. M. & Usov, V. V. *Astrophys. Space Sci.* **25,** 149 (1973).
11. Cavaliere, A. & Morrison, P. *Astrophys. J. Lett.* **238,** L63 (1980).
12. Neugebauer, G., Oke, J. B., Becklin, E. E. & Matthews, K. *Astrophys. J.* **230,** 79 (1979).
13. Richstone, D. & Schmidt, M. *Astrophys. J.* **235,** 361 (1980).
14. Ulrich, M. H. *et al. Mon. Not. R. astr. Soc.* **192,** 561 (1980).
15. Grandi, S. A. & Phillips, M. M. *Astrophys. J.* **239,** 475 (1980).
16. Netzer, H. & Davidson, K. *Mon. Not. R. astr. Soc.* **187,** 871 (1979).
17. Green, R. F. *et al. Astrophys. J.* **239,** 483 (1980).

An Antarctic iron meteorite contains preterrestrial impact-produced diamond and lonsdaleite

Roy S. Clarke Jr, Daniel E. Appleman & Daphne R. Ross

Department of Mineral Sciences, National Museum of Natural History, Smithsonian Institution, Washington DC 20560, USA

Of the many meteorites recovered so far from the Allan Hills, Antarctica, only nine have been irons. One of these, ALHA77283, contains troilite(FeS)–graphite(C)–schreibersite($(Fe,Ni)_3P$)–cohenite(Fe_3C) inclusions rich in the carbonado-type diamond–lonsdaleite 'nodules' previously described from the Canyon Diablo meteorite[1,2]. The Canyon Diablo, Arizona, meteorite, the excavator of Meteor Crater, is the only other iron meteorite known to contain these high-pressure minerals, and their occurrence in that meteorite has been explained as the result of shock-induced transformation of graphite, most probably at the moment of terrestrial impact and disintegration of the projectile during crater formation[3–6]. The suggestion[7] that formation was by high gravitational pressure has not been accepted. Virtually identical diamond–lonsdaleite-containing material in ALHA77283 occurs in a meteorite specimen with a well developed heat-altered zone produced by atmospheric ablation. It seems, therefore, that the diamond and lonsdaleite were present in the meteoroid before its final ablative passage through the atmosphere and soft landing on the ground. The shock event that produced these high pressure phases, therefore, must have taken place on its parent body or have been associated with the disruption of that body. Here we present the metallographic and X-ray diffraction data on which this conclusion is based.

Table 1 X-ray powder diffraction data for diamond–lonsdaleite nodules (CuKα radiation; Ni filter)

Lonsdaleite + cubic diamond*		Film 7123 Allan Hills A77283		Film 6990 Cayon Diablo USNM 2262	
d(Å)	I	d(Å)	I	d(Å)	I
2.18	4	2.19	M	2.17	M
2.06	10 D	2.09	S D	2.09	S D
1.933	2	1.94	MW	1.94	W
1.50	1				
1.257	6 D	1.26	M D	1.26	M D
1.17	1	1.18	VVF		
1.075	3 D	1.08	W D	1.07	W D

D, Reflections coincident with those of cubic diamond; S, strong; M, medium; MW, medium-weak; W, weak; VVF, very, very faint.
* From ref. 2.

ALHA77283 is a 10.5-kg carbon-rich octahedrite[8] of chemical group IA, containing 7.33% Ni, 69 p.p.m. Ga, 230 p.p.m. Ge, and 2.2 p.p.m. Ir. It is similar in composition and structure to the Canyon Diablo meteorite and distinct from the other iron meteorites so far recovered from the Allan Hills. It has a generally rounded anterior surface that is suggestive of aerodynamic shaping during oriented atmospheric flight, and a slightly concave posterior surface (Fig. 1). Regmaglypts are not a prominent feature although there are fluted areas on one side of the anterior surface near where it joins the posterior surface. Sectioning would be required to determine whether these are due to ablation or weathering. The specimen has been severely weathered, particularly on the posterior surface. Wind abrasion while on the ice also seems to have been important. Taenite (γ-FeNi, face-centred cubic) bands stand out in relief over much of the surface, revealing its internal octahedrite structure.

It was while cutting a median slice (Fig. 2) through the specimen that the presence of diamond was suspected. Cutting was slow, and the resulting surface revealed a cohenite-rich meteorite with troilite–graphite–schreibersite–cohenite inclusions containing black knobby protrusions up to ~0.5 mm in diameter, which protruded from narrow rims of black material of columnar structure that frequently separated troilite from schreibersite. These unusually hard inclusions prevented adequate polishing of the troilite inclusion and its immediate surroundings. Therefore, our observations on this association are tentative, but they do generally agree with the descriptions of the carbonado-type diamond–lonsdaleite 'nodules' previously reported in the Canyon Diablo meteorite[1–7].

Carbonado-like material was isolated from a small troilite inclusion on the back of the slice, area A in Fig. 2. The inclusion was dissolved in dilute hydrochloric acid, the insoluble residue treated several times with aqua regia, the light residue decanted from the heavy, and the heavy residue finally leached several times with boiling water. Part of this residue was examined by X-ray diffraction, revealing the presence of diamond, lonsdaleite, and a graphite-like phase. The diamond and lonsdaleite patterns are diffuse and the graphite pattern sharp. Diamond

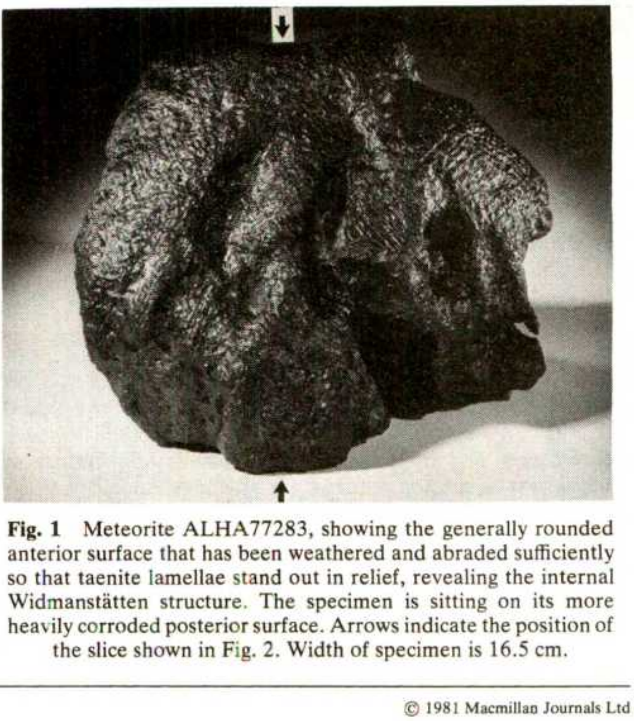


Fig. 1 Meteorite ALHA77283, showing the generally rounded anterior surface that has been weathered and abraded sufficiently so that taenite lamellae stand out in relief, revealing the internal Widmanstätten structure. The specimen is sitting on its more heavily corroded posterior surface. Arrows indicate the position of the slice shown in Fig. 2. Width of specimen is 16.5 cm.

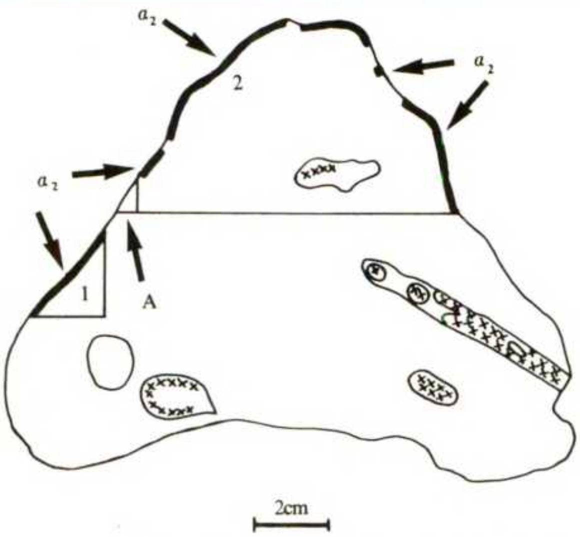


Fig. 2 Outline of slice through ALHA77283 perpendicular to the face shown in Fig. 1. Enclosed areas indicate troilite inclusions. ×, Distribution of graphite + diamond + lonsdaleite material identified visually on the rough cut surface. Areas 1 and 2 have been examined metallographically. Heavy black lines indicate presence of α_2 structure produced by ablative heating during atmospheric passage. The back side of area A contained a small troilite inclusion from which the material used for X-ray diffraction studies was obtained. The left edge in the sketch is the surface shown in Fig. 1.

and lonsdaleite are both present in appreciable amounts and are extremely fine-grained, with a grain size of the order of 100 Å. Preferred orientation is present in all three phases. Similar material was extracted from two previously-studied Canyon Diablo samples, USNM-1724 (refs 1, 4–6) and USNM-2262 (ref. 6). Both of these samples contained virtually identical assemblage of extremely fine-grained diamond, extremely fine-grained lonsdaleite, and coarser-grained graphite, with appreciable preferred orientation. These observations agree with previously reported results and are consistent with direct conversion of graphite to diamond and lonsdaleite by shock (Table 1).

Metallographic examination of polished sections (areas 1 and 2, Fig. 2) indicate that the specimen has been shock-loaded to moderate intensity[9–11] and that the shock event(s) took place preterrestrially. The following observations were made away from the heat-altered zone at the edge of the specimen. The kamacite (α-FeNi, body-centred cubic) matrix is moderately clear, containing slightly distorted Neumann bands, and patchy ε transformation structure that is frequently associated with cohenite grains (Fig. 3*a*). Taenite and plessite are present and appear unaffected. The structure is dominated by cohenite that is cracked, frequently offset, and badly plucked. The cohenite appears to be single crystal with a suggestion of undulatory extinction, and neither decomposition to graphite nor diffusion borders in kamacite were seen. Grain-boundary schreibersite is present, as are small schreibersites associated with cohenites. Rhabdites were not observed. Several small schreibersites at cohenite borders seem to have been partially melted, and small areas of adjacent kamacite converted to martensite. Some of these melted schreibersites are associated with areas of transformation within the cohenite, but other areas of transformed cohenite are seen without schreibersite[12]. No troilite sufficiently well polished was available for examination. Weathering within the structure is light and generally associated with cracked cohenite or schreibersite inclusions. These observations suggest shock loading in the moderate range of Heymann *et al.*[5], well above 200 kbar but less than the 750 kbar required to transform kamacite to martensite. The two specimens of the Canyon Diablo material we used for X-ray examination were also shocked in the moderate range[6]. These studies are continuing and will be published in detail elsewhere.

Metallographic examination revealed a well developed heat-altered zone over most of the available sample edge (Fig. 2), a structure that forms during ablative passage through the atmosphere[10,11]. In this zone the kamacite has been transformed to martensite (α_2, Fig. 3*b*), taenite has been rendered clear, schreibersite melted to a eutectic, and carbon diffusion zones produced around cohenite. Terrestrial oxides subsequently formed at the outer edge. The martensite zone is up to 1.5 mm thick, and where it is thin or absent it seems to have been lost to weathering. No remnant fusion crust was observed.

The presence of a heat-altered zone on ALHA77283 demonstrates that it entered the atmosphere from space, was ablatively decelerated, and landed softly. Therefore, the shock loading that produced the observed metallographic changes and the diamond and lonsdaleite must have taken place before atmospheric entry, but after Widmanstätten pattern formation. The most reasonable assumption is that these features were produced in the meteoroid at the time of parent body breakup or on subsequent collisions in space.

The close similarity between ALHA77283 and diamond–lonsdaleite-bearing Canyon Diablo specimens is striking.

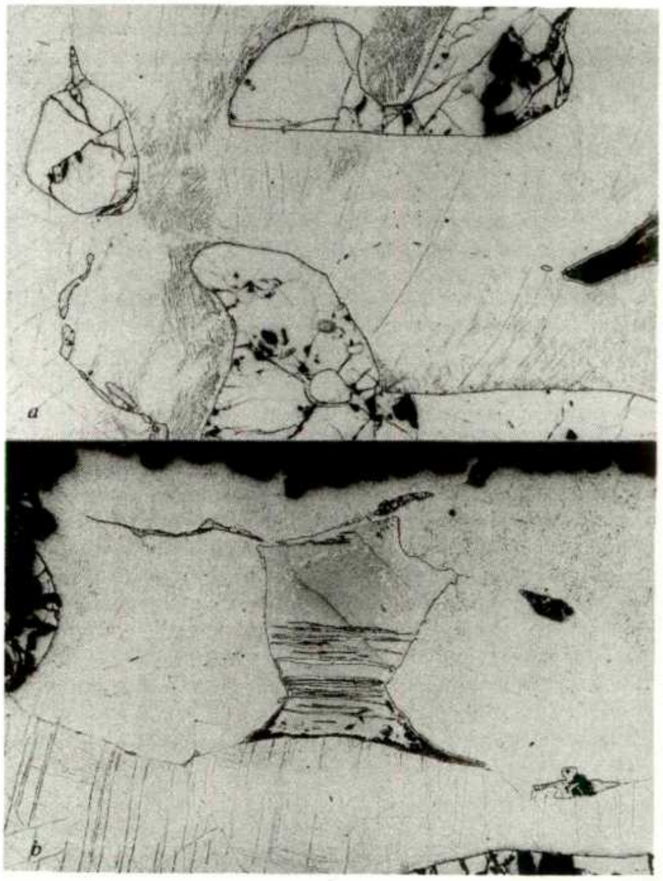


Fig. 3 Photomicrographs of metallographic structures revealed by acid etching of polished section 2 of meteorite ALHA77283. *a*, A 2.2-mm wide area at least 2 cm away from the nearest edge. The three large cracked inclusions are cohenite, the string of small inclusions at the lower left are schreibersite, and the dark inclusion at the middle right is a taenite-plessite area. The matrix is kamacite, containing distorted Neumann bands and patches of feathery ε transformation structure, particularly associated with edges of cohenite grains. The presence of the ε transformation structure establishes shock loading above 200 kbar. *b*, A 3.5-mm wide area of structure at the edge of the specimen. The ablation-produced heat-altered zone is at the top and is ~1 mm wide. The kamacite has been converted to granular martensite (α_2), the string of schreibersites at the top have undergone incipient melting at their edges, and the top of the taenite–plessite area does not colour when etched. In the less intensely heated area at the bottom, Neumann bands have been retained in the kamacite and the taenite-plessite colours on etching. These features are the consequence of a brief and steep thermal gradient produced by ablative heating.

Meteoroids have complex histories and shock events happen at various stages in their development. In certain circumstances sequential opportunities would be provided for shock to act on susceptible mineral associations to produce these diamond assemblages. This has been postulated as having taken place at the time of terrestrial impact for the canyon Diablo meteorite[4–6]. In the ALHA77283 meteoroid, however, diamond formation almost certainly took place preterrestrially. That this also happened in the Canyon Diablo meteoroid cannot be ruled out.

We thank Edward Anders and Michael E. Lipschutz for helpful discussions.

Received 30 January; accepted 19 March 1981.

1. Ksanda, C. J. & Henderson, E. P. *Am. Miner.* **24,** 677–680 (1939).
2. Frondel, C. & Marvin, U. B. *Nature* **214,** 587–589 (1967).
3. Nininger, H. H. *Arizona's Meteorite Crater* (American Meteorite Museum, Arizona, 1956).
4. Lipschutz, M. E. & Anders, E. *Geochim. cosmochim. Acta* **24,** 83–105 (1961).
5. Heymann, D., Lipschutz, M. E., Nielsen, B. & Anders, E. *J. geophys. Res.* **71,** 619–641 (1966).
6. Anders, E & Lipschutz, M. E. *J. geophys. Res.* **71,** 643–661 (1966).
7. Carter, N. L. & Kennedy, G. C. *J. geophys. Res.* **69,** 2403–2421 (1964).
8. Clarke, R. S. Jr., Jarosewich, E., Goldstein, J. I. & Baedecker, P. A. *Meteoritics* **15,** 273–274 (1980).
9. *Jain, A. V. & M. E. Lipschutz. Chem. Erde.* **30,** 199–215 (1971).
10. Axon, H. J. & Couper, W. R. D. *Mineral. Mag.* **40,** 827–841 (1976).
11. Buchwald, V. F. *Handbook of Iron Meteorites* (University of California Press, 1975).
12. Axon, H. J., Couper, W. R. D. & Kinder, J. *Nature* **267,** 414–415 (1977).

Excitation of the green line in the night airglow

R. D. Kenner, E. A. Ogryzlo & P. T. Wassell

Department of Chemistry, University of British Columbia,
Vancouver, British Columbia, Canada V6T 1Y6

In directions away from the Milky Way, most of the visible light from the night sky comes from a glowing layer of gas at an altitude of ~97 km. About 20% of this light is concentrated in a sharp green line at 557.7 nm which is a prominant feature of the visible portion of the night airglow. This emission comes from an excited oxygen atom ($O(^1S)$), but the mechanism by which this state is populated has been the subject of controversy ever since the emission was first identified in 1925 [1]. We present here new laboratory evidence concerning the processes governing the 557.7-nm emission of $O(^1S)$.

Because the maximum emission intensity of $O(^1S)$ occurs at an altitude where the oxygen atom concentration is at a maximum, it was assumed between 1931 and 1961 that the excitation occurs in the reaction[2]:

$$O + O + O \xrightarrow{k_1} O_2 + O(^1S) \qquad (1)$$

which has been called the Chapman reaction. To explain the absence of any green line emission in a laboratory system containing oxygen atoms, Barth[3] proposed an indirect mechanism:

$$O + O + M \xrightarrow{k_2} O_2{}^* + M \qquad (2)$$

$$O_2{}^* + O(^3P) \xrightarrow{k_3} O_2 + O(^1S) \qquad (3)$$

called the Barth mechanism. Laboratory studies have not yet provided clear evidence for or against either mechanism, although rocket measurements on the 557.7-nm intensity profile[4–6] apparently favour the Barth mechanism, and the theoretical considerations of Bates[7] suggest that the Chapman mechanism would be less efficient at producing $O(^1S)$. One complicating feature has been the magnitude of the quenching constant for the reaction:

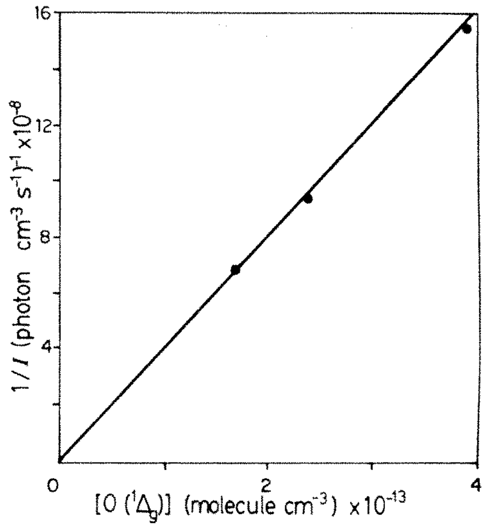


Fig. 1 The $O_2(^1\Delta_g)$ dependence of the 557.7-nm emission intensity (I). $[Ar] = 1.3 \times 10^{17}$ atom cm^{-3}, $[O] = 1.2 \times 10^{15}$ atom cm^{-3}.

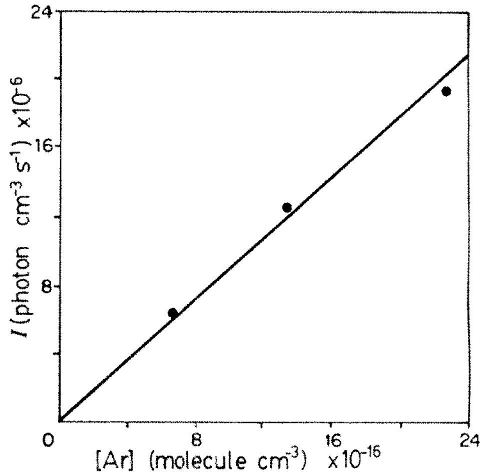


Fig. 2 The pressure dependence of the 557.7-nm emission intensity (I) over the range of 2–6 torr Ar. $[O_2(^1\Delta_g)] = 2 \times 10^{13}$ molecule cm^{-3}, $[O] = 7 \times 10^{14}$ atom cm^{-3}.

$$O(^1S) + O \xrightarrow{k_4} O(^1D) + O \qquad (4)$$

Reported values for k_4 have risen dramatically over the years[8]: 2.7×10^{-15} (1961), 1.8×10^{-13} (1966), 7.5×10^{-12} (1972) and 2×10^{-11} (1976) (in units of cm^3 per molecule per s). However, Slanger has found (personal communication) that the later high values are erroneous, with the error arising because of the presence of another quencher in these systems which is most probably $O_2(^1\Delta_g)$. The data presented here support this suggestion.

We prepared a stream of oxygen atoms by discharging a mixture of oxygen and argon, and passing the Ar–O_2–O mixture into a 5-l bulb. The oxygen atom concentration was obtained by monitoring the visible NO_2 emission when known quantities of NO were added[9]. The $O_2(^1\Delta_g)$ concentration was calculated from the emission intensity at 1,270 nm using a McPherson 0.3-m monochromator and a dry ice-cooled lead sulphide detector. The 557.7-nm emission was obtained by scanning the peak with a 1-m Spex monochromator.

The results of our measurements of the intensity (I) of the 557.7-nm emission as a function of $O_2(^1\Delta_g)$, Ar and O are contained in Figs 1–3. These show an inverse dependence on the $O_2(^1\Delta_g)$ concentration, a first-order dependence on the total particle concentration (principally Ar in this case) and a second-order dependence on the oxygen atom concentration. These experimental results can be summarized in the equation:

$$I(557.7\ \mathrm{nm}) = \frac{k[\mathrm{O}]^2[\mathrm{M}]}{[\mathrm{O_2}(^1\Delta_g)]} \qquad (5)$$

with the composite rate constant $k = 2.7 \times 10^{-27}$ cm^3 per molecule per s.

The rate law represented by equation (5) is not consistent with the simple Chapman reaction (equation (1)) but would result from the following 'Barth'-type mechanism, that is, formation by equations (2) and (3) followed by:

$$\mathrm{O_2}^* + \mathrm{O} \xrightarrow{k_3'} \text{quenched products} \qquad (3')$$

$$\mathrm{O}(^1\mathrm{S}) + \mathrm{O_2}(^1\Delta_g) \xrightarrow{k_6} \text{quenched products} \qquad (6)$$

$$\mathrm{O}(^1\mathrm{S}) \xrightarrow{k_7} \mathrm{O}(^1\mathrm{D}) + h\nu(557.7\ \mathrm{nm}) \qquad (7)$$

With standard steady-state assumptions, the composite rate constant in equation (5) is given by:

$$k = \frac{k_2 k_3 k_7}{k_3' k_6}$$

Although the rate of reaction (6) cannot be determined independently in our system, the relative efficiencies of O_2 and $O_2(^1\Delta_g)$ can be determined by adding O_2 to the gas stream. In the analysis we must add the reaction:

$$\mathrm{O}(^1\mathrm{S}) + \mathrm{O_2} \xrightarrow{k_9} \mathrm{O} + \mathrm{O_2} \qquad (9)$$

to the above reaction scheme. With standard steady-state assumptions for O_2^* and $O(^1S)$ the rate equation can be written:

$$\frac{[\mathrm{O}]^2[\mathrm{M}]}{I(557.7\ \mathrm{nm})[\mathrm{O_2}(^1\Delta_g)]} = \frac{k_6 k_3'}{k_3 k_2 k_7} + \frac{k_3' k_9 [\mathrm{O_2}]}{k_3 k_2 k_7 [\mathrm{O_2}(^1\Delta_g)]} \qquad (10)$$

A plot of the left-hand side of equation (10) against $[O_2]/[O_2(^1\Delta_g)]$ should yield a straight line with intercept/slope = k_6/k_9. Such a plot (with [O] and [M] constant) is shown in Fig. 4, from which $k_6/k_9 = 2.5 \times 10^3$. As $k_9(298\ \mathrm{K}) = 2.2 \times 10^{-13}$ cm^3 per molecule per s (ref. 10) our results yield $k_6(298\ \mathrm{K}) = 5.5 \times 10^{-10}$ cm^3 per molecule per s. This very large rate constant is comparable with that reported for the quenching of $O(^1S)$ by O_3 (ref. 11).

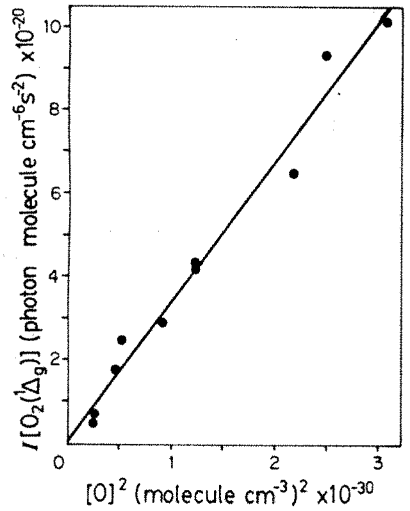


Fig. 3 The atomic oxygen dependence of the 557.7-nm emission intensity (I), corrected for different $[O_2(^1\Delta_g)]$. $[\mathrm{Ar}] = 1.3 \times 10^{17}$ atom cm^{-3}.

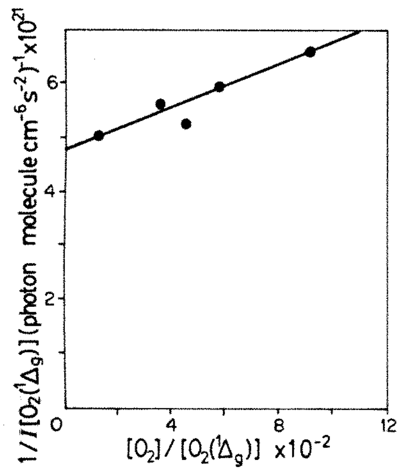


Fig. 4 The effect of the addition of O_2 on the 557.7-nm emission intensity as a function of $[O_2(^1\Delta_g)]$. $[\mathrm{O}] = 7 \times 10^{14}$ atom cm^{-3}, $[\mathrm{Ar}] = 1.3 \times 10^{17}$ atom cm^{-3}.

Although quenching of $O(^1S)$ by $O_2(^1\Delta_g)$ clearly dominates the 557.7-nm emission in the laboratory, it is difficult to see it playing a significant part in the upper atmosphere. From the 60 k R emission from $O_2(a^1\Delta_g)$ in the 97-km region one can estimate a $O_2(a^1\Delta_g)$ concentration of $\sim 1.8 \times 10^8$ molecule cm^{-3}. Combining this with k_6 yields a removal rate an order of magnitude slower than the radiative decay rate of $O(^1S)$, which is 1.065 s^{-1}.

We conclude that the recombination of oxygen atoms leads to the excitation of $O(^1S)$ through a Barth-type mechanism in which the intermediate O_2^* remains to be identified. The quenching of $O(^1S)$ by $O_2(^1\Delta_g)$ occurs at collision frequency and probably represents a very efficient energy-pooling process[12] which results in the dissociation of $O_2(^1\Delta_g)$.

This work was supported by the Air Force Office of Scientific Research, Air Force Systems Command USAF under the grant AFOSR 79-0088.

Received 11 March; accepted 25 March 1981.

1. McLennan J. C. & Shrum, G. M. *Proc. R. Soc.* **A108,** 501–512 (1925).
2. Chapman, S. *Proc. R. Soc.* **A132,** 353–374 (1931).
3. Barth, C. A. *J. geophys. Res.* **67,** 1628 (1961).
4. Slanger, T. G. & Black, G. *Planet Space Sci.* **25,** 79–88 (1977).
5. Thomas, L., Greer, R. G. H. & Dickinson, P. H. G. *Planet Space. Sci.* **27,** 925–931 (1979).
6. Witt, G., Stegman, J. Solheim, B. H. & Llewellyn, E. J. *Planet. Space Sci.* **27,** 341–350 (1979).
7. Bates, D. R. *Planet. Space Sci.* **27,** 717–718 (1979).
8. Bates, D. R. *Planet. Space Sci.* **26,** 897–912 (1978).
9. Sutoh, M., Morioka Y. & Nakamura, M. *J. chem. Phys.* **72,** 20–24 (1980).
10. Slanger, T. G., Wood, B. J. & Black, G. *Chem. phys. Lett.* **17,** 401–403 (1972).
11. Atkinson, R. & Welge, K. H. *J. chem. Phys.* **57,** 3689–3693 (1972).
12. Arnold, S. J., Finlayson, N. & Ogryzlo, E. A. *J. chem. Phys.* **44,** 2529–2530 (1966).

Porous titania glass as a photocatalyst for hydrogen production from water

Leonard Kruczynski, Hyman D. Gesser, C. William Turner & Edward A. Speers*

Department of Chemistry, University of Manitoba, Winnipeg, Canada R3T 2N2

The rutile modification of TiO_2 in the form of a powder suspended in water shows promise as a photocatalyst for the production of hydrogen from water[1,2]. Although hydrogen production was reported[3] from pure rutile powder suspended in an aqueous solution, subsequent work[1,2] revealed that sustained hydrogen production required Pt or RuO_2 on the rutile surface. This limiting usefulness of crystalline rutile in photocatalytic applications has prompted us to investigate the photocatalytic properties of amorphous titania in the form of a porous glass. Here we report an inexpensive method of forming monolithic porous titania glass (PTG) and some properties that make it uniquely suited to photocatalytic applications such as its inexpensive fabrication, its high surface area, and its spontaneous photoformation of H_2 from water without the need for precious metals.

---

* Present address: Industrial Research and Development, 345 Higgins Avenue, Winnipeg, Manitoba, Canada R3A OV5.

0028-0836/81/220399—03$01.00

Table 1 Comparison of hydrogen production rates from PTG with those obtained by other methods

Catalyst and solution		Form of catalyst	Amount of catalyst (g)	Rate of H_2 production (μmol h^{-1})	Ref.
$SrTiO_3$	Conc. NaOH	Single crystal	—	0.1	17
TiO_2	Sugar	Powder	0.3	0.01	18
$TiO_2/RuO_2/Pt$	Sugar		0.3	14.0	18
$TiO_2/RuO_2/Pt$	H_2O		0.3	0.02	18
$TiO_2/RuO_2/Pt$	Sugar + 6 M NaOH		0.3	17.0	18
TiO_2	50% MeOH		0.3	27.0	19
$TiO_2/RuO_2/Pt$	50% MeOH		0.3	520.0	19
PTG	H_2O	Glass‡	1.5	~0.01	This work
PTG	50% KOH		1.5	0.03	
PTG	MeOH + 50% KOH		1.5	6.0	
PTG/Cu	MeOH + 50% KOH		1.5	27.0	
PTGB*	50% KOH		1.5	Trace	
PTGB*	MeOH + 50% KOH		1.5	14.0	
PTGB/Cu	MeOH + 50% KOH		1.5	280	
Translucent aerogel†	H_2O	Chunks	1.5	912	

*PTGB, porous titania glass treated treated with base: soaked in 50% KOH, dried at 80 °C overnight.

†TiO_2 produced by hydrolysing $Ti(OiPr)_4$ and autoclaving the gel at 270 °C.

‡Particles averaging 5 mm in diameter, 1 mm thick.

Although titania gels were first reported in 1823[4], no successful attempt to dehydrate the gel to a single piece of porous glass has been reported. Titania gels were originally formed using inorganic methods: adding $TiCl_4$ to concentrated HCl followed by partial neutralization[5], or adding acid to a sodium titanate solution[6]. Inorganic methods produce salts which must be removed by dialysis before the sol–gel transition will occur[6]. In an effort to eliminate the presence of salts and hence the need for dialysis, organo-titanium reagents have been used to form titania gels. The hydrolysis of titanium alkoxides results in the formation of amorphous powders with large surface areas, but not monolithic (pieces larger than 1 cm^2) samples[7]. More recently, monolithic glasses containing titania have been reported[8]. These mixed titania–silica glasses are formed by polymerizing a mixture of silicon and titanium alkoxides. In view of the low cost of $TiCl_4$, we modified an inorganic method of producing PTG in the form of flat wafers 0.5–1 mm thick and up to 3.5 cm in diameter[9].

The dropwise addition of 6 mol of water to 1 mol of $TiCl_4$ at low temperatures results in a clear, colourless solution. Up to 60% of the Cl^- in this solution can be removed as HCl (corresponding to 2.4 mol of HCl per mol of $TiCl_4$) by low pressure, room temperature distillation. The viscous residue can be redissolved in water. Careful addition of KOH to this solution results in the neutralization of 0.2 mol HCl; adding more KOH results in permanent turbidity. From this partially neutralized solution, a stock solution containing 0.35 mol Ti/L, can be kept for months at 4 °C, but becomes turbid on standing at room temperature for a few days. Dialysis of the stock solution results in a rapid increase in viscosity after 3–4 h; in 12 h the dialysed solution sets to a clear gel. After drying for 10 days at room temperature, the gel contracts to 30% of its original volume. Dehydration of this air-dried gel at 160 °C results in a weight loss of 22%, but no change in size; this high temperature dehydration gives a microporous structure. Additional heating above 160 °C results in a further 2% weight loss, followed by the onset of crystallization to anatase at 350 °C. A similar crystalline phase change occurring at 300 °C has been reported for a precipitated amorphous titania powder[10].

The IR spectrum of PTG outgassed at 220 °C (0.01 torr) for 10 h is shown in Fig. 1*a*. Sharp bands in the 3,700 cm^{-1} region are assigned to surface O–H stretching frequencies unperturbed by hydrogen bonding[11]. In this region PTG shows two bands, at 3,659 and 3,638 cm^{-1}; quite different from either rutile (3,680 cm^{-1})[12] or anatase (3,715, 3,675 cm^{-1})[13]. Heating the glass to higher temperatures results in the loss of both bands due to surface dehydroxylation (Fig. 1 *a–c*). Deuteration of the surface by exposure to D_2O liquid followed by outgassing at 330 °C (Fig. 1*d*) gives a ν_{OH}/ν_{OD} ratio 1.36 compared with the expected 1.37 isotope shift for an isolated OH group. The BET surface area decreases from 490 $m^2\ g^{-1}$ for a sample heated at 300 °C to 146 $m^2\ g^{-1}$ for samples heated at 350 °C. The crystallization to anatase (as determined by X-ray powder patterns) that occurs in the 300–350 °C range not only accounts for the decrease in surface area, but also renders the glass visibly opaque. Surface area changes, weight loss, and crystallization behaviour determined for PTG are all very similar to amorphous titania prepared by precipation from solution by base addition[14]. PTG dehydrated at temperatures <350 °C has a pale yellow colour—this colour and ESR data indicate the presence of surface Ti(IV) peroxo species. The ESR signals observed on PTG are the same as those observed in anatase treated with H_2O_2 and assigned to O_2^- species[15].

Rutile produces hydrogen photocatalytically from water when exposed to visible light only when the rutile surface also contains Pt or RuO_2 (refs. 1,2,16). This can be explained by the close proximity of the conduction band edge of rutile to the reduction potential of water (Fig. 2), necessitating the presence of precious metals to reduce the overpotential for H_2 evolution at the surface. On the other hand, the upward displacement of the conduction band edge in $SrTiO_3$ above the reduction potential of water (Fig. 2) may explain why precious metals are not necessary for photocatalytic hydrogen production in $SrTiO_3$ (ref. 17). As PTG, like $SrTiO_3$, will produce H_2 from water without the intervention of precious metals, we conclude the PTG must have a bandgap structure similar to $SrTiO_3$. Furthermore, $SrTiO_3$ and the anatase modification of TiO_2 have identical bandgaps on 3.2 eV. From X-ray powder studies, PTG closely resembles anatase, being transformed into anatase at 350 °C.

When untreated PTG is immersed in water and illuminated with visible light from a 2,500 W high pressure Xenon lamp operated at 65% of rated power to match the solar spectrum in frequency distribution and intensity, small amounts of H_2 are produced as measured by gas chromatography (2 m, 5 A molecular sieve column, Argon carrier gas with thermal conductivity detector). When PTG is immersed in 50% KOH and irradiated, it turns intense blue-black in colour and the rate of H_2 production increases slightly. The blue-black colour is due to the reduction of Ti (IV) to blue-black Ti(III). The intense colour is

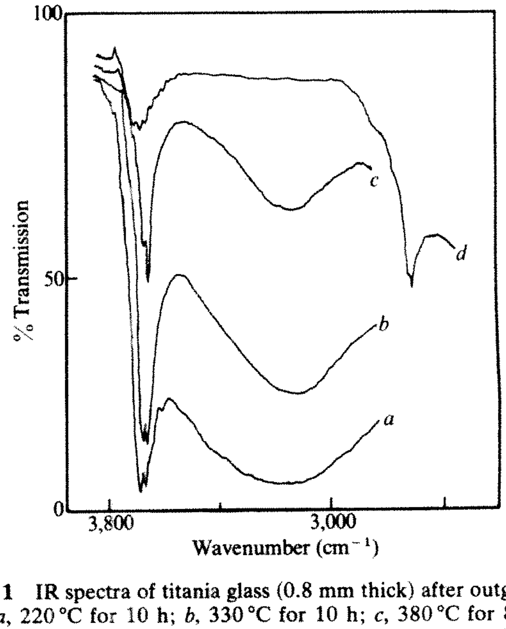


Fig. 1 IR spectra of titania glass (0.8 mm thick) after outgassing at: *a*, 220 °C for 10 h; *b*, 330 °C for 10 h; *c*, 380 °C for 8 h; *d*, exchange with D_2O (liquid) followed by 8 h outgassing at 330 °C.

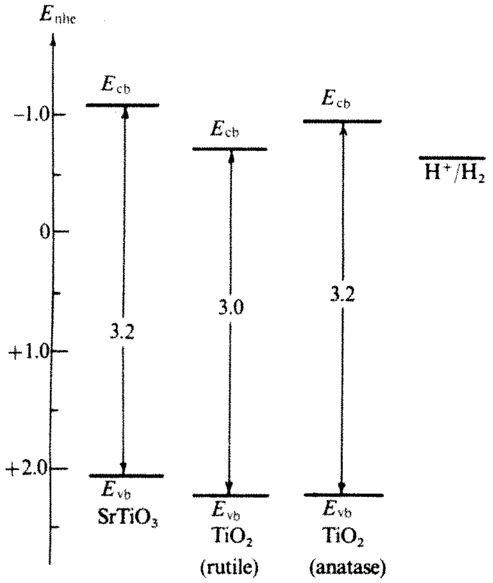


Fig. 2 A comparison of bandgap energy positions in anatase, rutile and $SrTiO_3$ relative to the reduction potential of water (H^+/H_2). Energy levels are shown for pH 7.

stable over several days as long as there is no contact with O_2, and will remain for about a week in the air if totally immersed in solution. By providing an oxidizable substance such as methanol, the colour will disappear more rapidly and H_2 production will increase dramatically. Irradiation experiments with CH_3OH in D_2O produced D_2 and not H_2; with added CD_3OH in H_2O only H_2 was formed. Thus H_2 produced in this way comes from water and not the alcohol. A further addition of 0.3 mol of Cu(II) in the form of $CuCl_2$ solution increases the rate of H_2 production further. These results are summarized and compared with previous work in Table 1.

If PTG is soaked in 50% KOH and then dried before exposing to light, photocatalytic H_2 production is doubled with respect to untreated PTG. The addition of the Cu(II) solution to PTGB (PTG base treated) increases the rate of H_2 evolution by an order of magnitude. The mechanism of Cu enhancement is not entirely clear. The Cu(II) is removed from solution, probably becoming incorporated into the TiO_2 structure. The PTG or PTGB becomes a purplish colour on exposure to light, indicating that copper is in the form of the free metal or Cu_2O; the purple colour resulting from the combination of the red of Cu metal or Cu(I) oxide and the blue-black of Ti(III). The Cu-treated glasses revert to a pale greenish yellow on exposure to air.

The data summarized in Table 1 show some trends worth noting: the rate of H_2 production can be increased by the presence of oxidizable substances such as sugar or methanol and/or the presence of strong base such as NaOH or KOH. As the catalysts are of different forms, a direct comparison may not be valid as the mechanisms for the H_2 production may depend on the form of the catalyst, surface area and light intensity.

We thank the National Research Council of Canada for support.

Received 2 January; accepted 17 March 1981.

1. Sato, S. & White, J. M. *Chem. Phys. Lett.* **72,** 83–86 (1980).
2. Kawai, T. & Sakata, T. *Chem. Phys. Lett.* **72,** 87–89 (1980).
3. Schrauzer, G. M. & Guth, T. D. *J. Am. chem. Soc.* **99,** 7189–7193 (1977).
4. Rose, H. *Liebigs Ann.* **73,** 76 (1923).
5. Weiser, H. B. *J. phys. Chem.* **46,** 1051–1059 (1942).
6. Klosky, S. & Marzano C. *J. phys. Chem.* **29,** 1125–1128 (1925).
7. Harris, M. R. & Whitaker G. *J. appl. Chem.* **12,** 490–494 (1962).
8. Yoldas, B. E. *J. Mater. Sci.* **14,** 1843–1849 (1979).
9. US Patent Pending.
10. Vivien, D., Levage, J. & Mazieres, C. *J. chim. Phys.* **67,** 199–204 (1979).
11. Yates, D. J. *J. phys. Chem.* **65,** 746–753 (1961).
12. Jackson, P. & Parfitt, G. D. *Trans. Faraday Soc.* **67,** 2469–2483 (1971).
13. Primet, M., Pichat, P. & Mathieu, M-V. *J. phys. Chem.* **75,** 1216–1220 (1971).
14. El-Akkad, T. M. *J. Coll. Interface Sci.* **76,** 67–73 (1980).
15. Munuera, G., Gonales-elipe, A. R., Soria, J. & Soz, J. *JCS Faraday I.* **76,** 1535–1546 (1980).
16. Borgarello, E., Kiwi, J., Pelizzette, E., Visca, M. & Grätzel, M. *Nature* **289,** 158–159 (1981).
17. Wagner, F. T. & Somarjai, G. A. *Nature* **285,** 559–560 (1980); *J. Am. chem. Soc.* **102,** 5494–5502 (1980).
18. Kawai, T. & Sakata T. *Nature* **286,** 474–476 (1980).
19. Kawai, T. & Sakata, T. *JCS Chem. Commun.* 694–695 (1980).

Scanning electron-stimulated desorption microscopy of model-biological surfaces

H. F. Dylla*, J. H. Abrams†, C. T. Hovland‡ & J. G. King§

* Plasma Physics Laboratory, Princeton University, Princeton, New Jersey 08544, USA
† Department of Surgery, University of Minnesota, Minneapolis, Minnesota 55455, USA
‡ Perkin-Elmer/Physical Electronics Division, Eden Prairie, Minnesota 55344, USA
§ Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Electron-stimulated desorption (ESD) of ions and neutrals from surfaces has been used to study chemisorption in many simple gas–solid systems[1,2], but there has been much less effort devoted to ESD studies with significant spatial resolution[3–7]. ESD contrasts with scanning electron microscopy because ESD signals depend sensitively on the coverage and binding of an adsorbed species[1]. Applying scanning ESD to biological surfaces permits the surface chemical variations to be mapped using simple low mass adsorbates[8] in comparison with conventional electron microscopy of biological matters which relies on heavy-metal stain technology[9]. We show here how scanning ESD can complement scanning Auger electron microscopy in mapping the surface chemistry of biological substrates. A model biological substrate constructed from a deposited synthetic polypeptide monolayer (polymethyl-glutamate) is used as the controlled sample surface. Micrographs display a spatial variation in H^+ desorption with a resolution of 10^{-4} cm. We also discuss limitations to resolution as dictated by signal-to-noise considerations. The contrast mechanism would be useful for studying the important biological surface problems such as membrane morphology, cell-interfacial composition, and cell-to-cell adhesion.

The micrographs were made using a scanning Auger microprobe[10]. A scanning electron gun, consisting of a LaB_6 source, electrostatic lenses and deflection plates, is concentrically mounted within a cylindrical mirror electron energy analyser. The electron gun can project a bean with a diameter of 1,600 Å with currents of $10^{-8} - 10^{-7}$ A onto the target. With a defocused beam (diameter ~ 1 μm) larger currents (~ 10^{-6} A) can be projected and scanning Auger mocrographs (SAM) and standard secondary electron micrographs (SEM) can be obtained[10]. Ions desorbed by the electron beam striking the target are energy-analysed electrostatically[11] and then mass-analysed with a quadrouple mass filter before detection with a CuBe multiplier. The amplified output of this detector was used for intensity modulation of an oscilloscope to produce scanning desorption micrographs (SDM). The entire apparatus is housed in a stainless steel ultrahigh vacuum chamber pumped by a combination of ion and titanium sublimation pumps to a base pressure of the order of 10^{-10} torr (~ 10^{-8} Pa). The small signal counting rates encountered in the SAM and SDM modes required exposure times of 15–40 min.

The model biological surface consisted of a monolayer of poly-λ-methyl-glutamate (PMG) adsorbed onto a clean, polished stainless steel substrate. A pure carbon grid was subsequently vacuum deposited over the poly-amino acid layer. The sample was prepared as follows: The stainless steel sub-

0028-0836/81/220401—04$01.00

strates, ~1 cm in diameter, were machined from 304 stainless steel and polished to optical flatness with alumina and diamond abrasives. The substrates were cleaned by rinsing in a sulphuric–chromic acid solution followed by a trichloroethylene, acetone, ethanol, and distilled water rinses. Poly-λ-methyl-glutamate[12] (0.9 mg) was dissolved in reagent grade chloroform (2 ml). A film of PMG cast onto distilled H_2O contained in a Pyrex trough was then deposited onto the stainless steel substrates by the Langmuir–Blodgett technique[13]. The PMG films were dried for 12 h at room temperature in a dust free environment. A carbon grid was vacuum deposited onto the PMG surface by shadowing the carbon flux with a 400 mesh electron microscope grid. The sample was rinsed in distilled water immediately before mounting it onto the specimen stage of the scanning Auger microprobe. The sample remained in the vacuum system for 1 h before analysis to allow the system pressure to fall to the 10^{-9} torr range.

Figure 1 shows an SEM view of C/PMG sample, and three different scanning Auger micrographs (SAM) of the same sample area. The relative elemental surface distribution of O, C and N is also shown. The C matrix is visible with high contrast in the standard SEM of the sample because of the large difference in secondary emission between C and the substrate. The C-SAM verifies the larger surface concentration of C at the matrix positions. The standard SEM gives no information of the presence of the PMG monolayer with this spatial resolution; only the topographical defects are shown on the underlying stainless steel substrate. The O-SAM shows attenuated oxygen signal from the $FeO–Cr_2O_3$ oxide layer on the stainless steel substrate[14], whereas the N-SAM shows the PMG overlayer as a result of the surface amide concentration.

Single point Auger spectra which aided the interpretation of the SAMs are shown in Fig. 2 for the analysis beam centred over a pure carbon area (*a*), a PMG coated area (*b*), and cleaned substrate area (*c*) of the sample. The Auger spectra in the pure C region shows only C (at 272 eV) and a small O (510 eV) peak. The small relative surface concentration of O (<20 atomic %) after saturation of the sample with H_2O shows the essential hydrophobic nature of pure C at room temperature. The fact that the underlying high energy Fe Auger transitions (at 598, 651, 703 eV) are not visible indicated that the C film thickness is >50 Å.

Comparing the point Auger spectra from the PMG area with the cleaned substrate reveals the attenuation of the Fe and O transition by the PMG overlayer. In particular, the low energy Fe transition at 47 eV, for which escaping electrons have an attenuation length[15] of ~5 Å, is completely attenuated by the PMG film. Assuming a simple model for the attenuation of low energy electrons:

$$I/I_0 = \exp(-\lambda/d \sin\theta) \tag{1}$$

where λ = the inelastic attenuation length, d, the film thickness, and θ, the collection angle, the PMG film thickness is estimated to be of the order of 10–20 Å. Some fraction of the O Auger signal observed on the PMG is due to O in the PMG film, but this is difficult to estimate because of the presence of O in the stainless steel substrate. However, N is not present in significant quantities in the substrate, and we conclude that all the observed N in the PMG spectra and the N-SAM, is from PMG film. A rough calculation of the N concentrations based on relative sensitivity factors[16] gives ~10 atomic % compared with the expected 16 atomic % for pure PMG. Little electron radiation damage was noted in the form of changes in the Auger spectra for the beam energy (2 keV), currents (75 nA), and exposure times (100–300 s) necessary for producing the micrographs.

Figure 3 shows a scanning desorption micrograph (SDM) of the C/PMG sample imaged with desorbing H^+ ions. For comparison with the sample topography, an adsorbed current micrograph of the same idea is also shown in Fig. 3. The desorbing H^+ ions are seen only over the PMG area and not over the pure C areas. The detected H^+ desorption signal was ~300 H^+ ions per second or ~75 H^+ ions per raster grid point. The shot noise in such a small signal accounts for the spotty appearance of the SDM.

The source of H^+ is assumed to be ESD of adsorbed H_2O on the PMG although electron impact decomposition of the PMG layer may also contribute to the signal. The boundary between the PMG monolayer and adsorbed molecules is probably not distinct, and incident low energy electrons can interact with molecular groups in both phases. H^+ can originate from adsorbed H_2O or CH_3 ligands; similarly OH^+ or O^+ can be desorbed from adsorbed H_2O or OH ligands. No ESD of OH^+ or O^+ ions was detected from the PMG sample. Its absence is consistent with the small size of the detected H^+ signal. Studies of ESD of

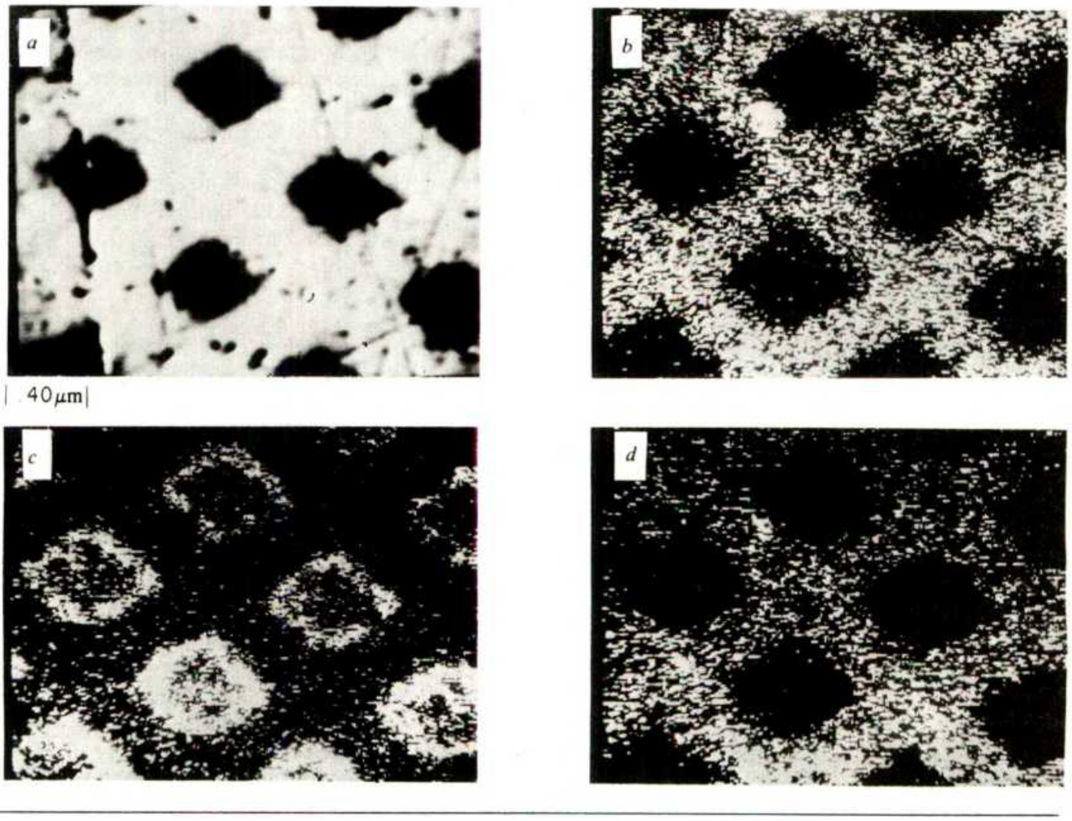


Fig. 1 Scanning electron micrographs of test sample consisting of mechanically polished 304 stainless steel, on which a monolayer of poly-λ-methyl-glutamate (PMG) was adsorbed. A rectangular array of carbon was deposited over the unmasked portion of the PMG by evaporation of carbon through a 400 mesh stainless steel grid. *a*, Standard scanning electron micrograph (SEM) of the sample taken with a 2 keV primary beam. The remaining images are scanning Auger micrographs taken with: *b*, O(510 eV); *c*, C(272 eV); and *d*, N(381 eV) Auger electron emission. The SAMs were taken with 2 keV, 75 nA, primary beam, which scanned 250 lines per frame at a frame time of 187 s for C and O and 1,250 s for N. Magnification ×400.

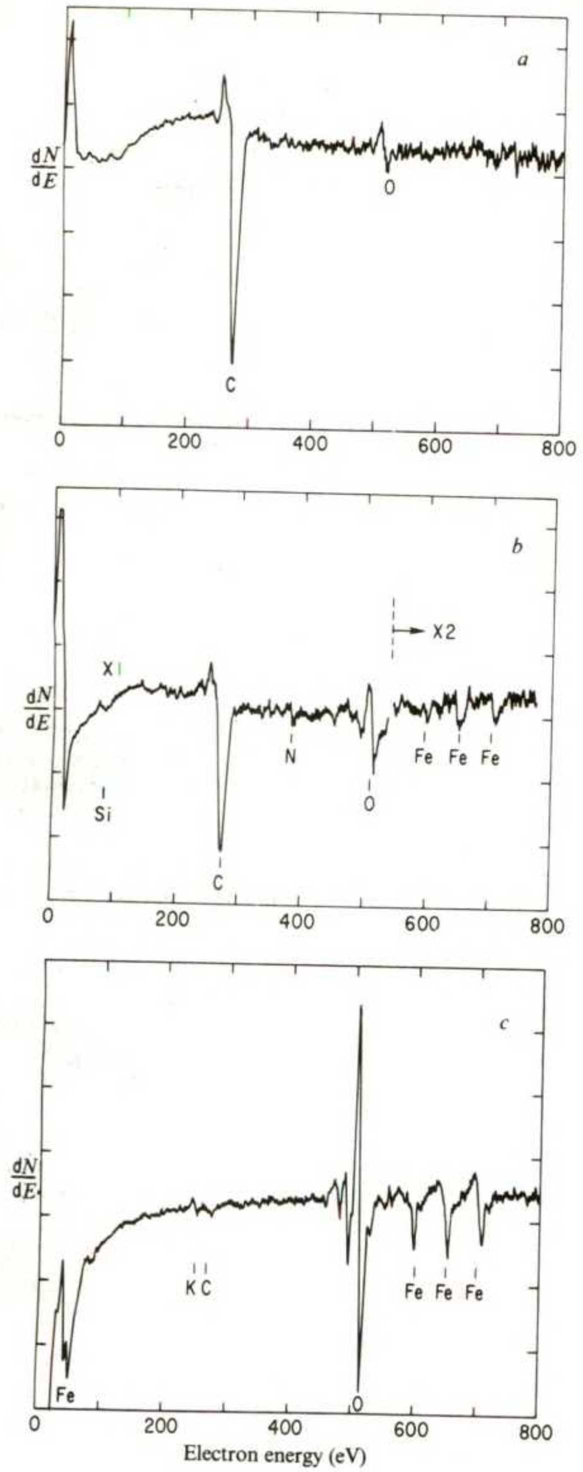


Fig. 2 Auger spectra of the C/PMG sample and substrate: spectra obtained when the primary bean is centred on: *a*, pure carbon area; *b*, PMG coated area of the sample; *c*, spectra obtained from the cleaned and degreased stainless steel substrate used for deposition of the C/PMG overlayer.

adsorbed H_2O on other surfaces[1,6] usually show H^+ signals at least an order of magnitude larger than OH^+ or O^+ desorption signals.

To check the sample preparation procedure, the sample was removed from the microprobe and the PMG overlayer was dissolved in $CHCl_3$ for inspection by mass spectrometry. The parent peak (at 147 AMU), and the expected fragments of the PMG monomer were clearly visible in the mass spectrum as the solution was heated to 70 °C.

The difference in the ESD of H^+ from a hydrophobic form of carbon (pure carbon) and a hydrophilic form of carbon (a proteinaceous hydrocarbon, PMG) has been used to image a matrix sample of the two materials exposed to H_2O. The relatively small desorption signal for this system (H^+ count rate of $\sim 300\ s^{-1}$) could be improved by incorporating an ion lens designed for high efficiency collection of ESD ions. The ion lens for this study was designed for SIMS for which overall signal count rates are rarely a constraint, but a small focusing area (and a small solid angle of acceptance) is the usual primary consideration[17]. The estimated ion collection efficiency for the ion lens system in this study is $\sim 10^{-3}$. This could be improved to $\sim 10^{-1}$ if the design constraints for SIMS were removed and the collection efficiency maximized. Any improvement in ion collection efficiency would allow shorter scan times per micrograph, and multiple adsoption–desorption cycles could be used in conjunction with the image storage and averaging to improve the signal-to-noise ratio.

The signal-to-noise ratio for SDM depends on the electron beam spot diameter; current density, J; scan time, Δt; ESD ionic cross-section, Q^+ and total cross-section, Q; and ion collection efficiency, ε. The dependence of the signal-to-noise ratio on the lateral resolution (d) can be easily calculated, assuming the primary source of noise will be the shot noise in the comparitively small detected particle count (S) during an interval Δt. For desorbed ion signals, the signal-to-noise ratio will be

$$\mathrm{S/N} \simeq S/S^{1/2} = \left(\varepsilon \frac{Q^+}{Q} n_0\right)^{1/2} d\left(1 - \exp\left(-\Delta t/\tau\right)\right)^{1/2} \qquad (2)$$

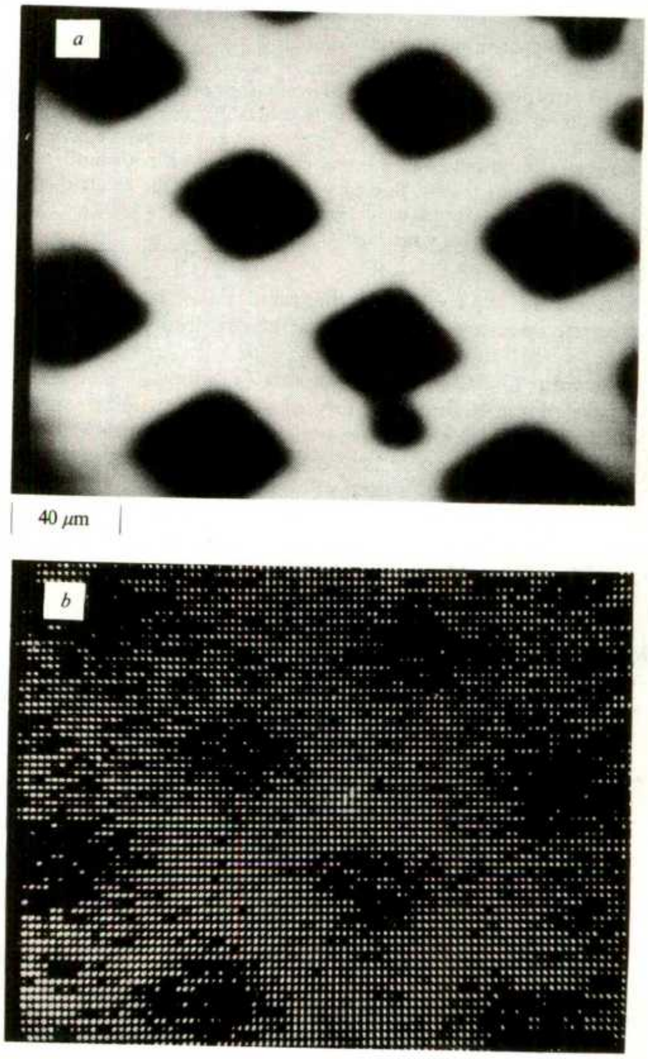


Fig. 3 *a*, Absorbed current micrograph of a neighbouring area of the sample shown in Fig. 1. *b*, Scanning desorption micrograph (SDM) of the same area imaged with desorbing H^+ ions. The beam voltage was 2.0 keV, beam current 1.7 µA, with 100 lines per frame scanned in 2,500 s. Magnification ×400.

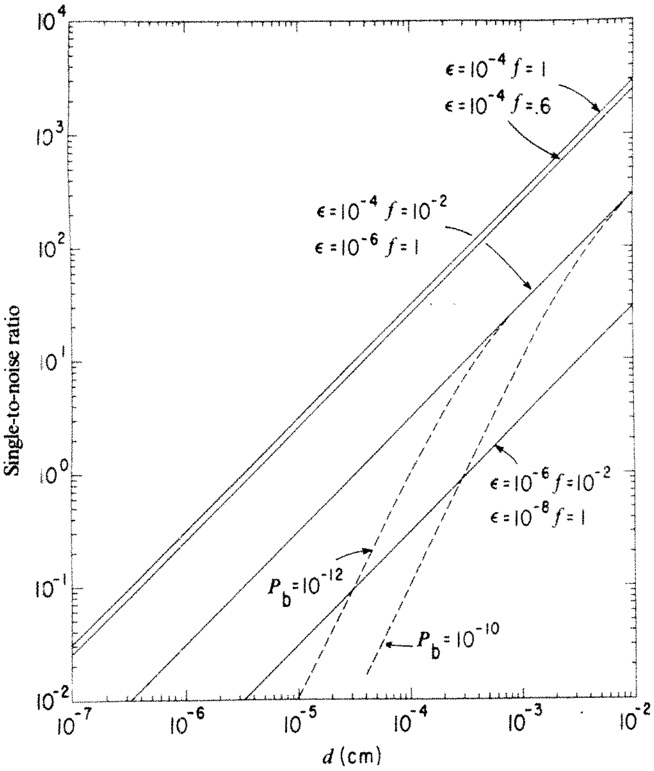


Fig. 4 Signal-to-noise ratio compared with resolution, where d = minimum resolvable diameter, for scanning electron-stimulated desorption microscopy where only shot noise is considered. The curves are plotted for various values of the total detection efficiency, ε, and the desorption fraction $f = (1 - \exp(-\Delta t/\tau))$, assuming a full monolayer initial coverage and a scan time per distance d, of $\Delta t = 2$ ms. Two curves are shown for the detection of desorbed (28 AMU) neutrals in the presence of a background pressure (10^{-10}–10^{-12} torr).

where $f \equiv (1 - \exp(-\Delta t/\tau))$ is the fraction desorbed of the initial adsorbate coverage n_0. The characteristic desorption time τ is e/JQ.

The ratio Q^+/Q is typically a small number[1], $\sim 10^{-2}$–10^{-4}, because of the high probability of Auger neutralization of a desorbed ion near a metal or semiconductor surface[18]. Therefore, one might expect larger ESD signals from desorbed neutrals where the full cross-section Q is applicable. However, this gain is compensated by the difficulty in detecting neutral species (mass spectrometer ionizer efficiencies[19] are at best 10^{-2}) and the additional noise source due to ionization of background gas molecules.

With this latter contribution the resulting signal-to-noise ratio for desorbed neutral signals is

$$\frac{\varepsilon n_0 d^2 (1 - \exp(-\Delta t/\tau))}{[\varepsilon n_0 d^2 (1 - \exp(-\Delta t/\tau)) + BP_B \Delta t]^{1/2}} \qquad (3)$$

where P_B is the background partial pressure of the molecule of interest and B is a constant related to the ionizer efficiency. Equations (2) and (3) are plotted in Fig. 4 for several values of the detection efficiency, ε, and desorption fraction, f. The straight lines in Fig. 4 correspond to desorbed ion signals or to neutral signals where the background pressure of the desorption species is negligible.

For instrumental parameters for the microprobe used in this study[10] and for ESD ionic cross-sections of 10^{-19} cm^2 for the adsorbate, the calculated spatial resolution limit is 2×10^{-4} cm for scanning desorption micrographs of minimal quality (signal-to-noise ratio is ~ 5).

Received 22 September 1980; accepted 16 March 1981.

1. Madey T. E. & Yates J. T. *J. Vac. Sci. Technol.* **8**, 525–555 (1971).
2. Menzel, D. *Surface Sci.* **47**, 370–383 (1975).
3. Rork, G. & Consoliver, R. F. *Surface Sci.* **10**, 291–294 (1968).
4. Lichtman, D. & Campuzano, J. *J. appl. Phys. Suppl.* **2**, 189–192 (1974).
5. Dylla, H. F. & King, J. G. *Bull. Am. phys. Soc.* **21**, 241 (1976).
6. Joshi, A. & Davis, L. E. *J. Vac. Sci. Technol.* **14**, 1310–1313 (1977).
7. Poppa, H. & Bauer, E. *Surface Sci.* **97**, L309–L312 (1980).
8. Weaver, J. C. & King, J. G. *Proc. natn. Acad. Sci. U.S.A.* **70**, 2781–2784 (1973).
9. DeMee, P. B. Frederickson, R. G. & Pope, R. S. in *Proc.* 1977 *Conf. on Scanning Electron Microscopy* Vol. II, 83-92 (IIT Research Institute, Chicago, 1977).
10. MacDonald, N. C., Hovland, C. T. & Gerlach, R. L. in *Proc.* 1977 *Conf. on Scanning Electron Microscopy.* Vol. I, 201–210 (IIT Research Institute, Chicago, 1977)
11. Gerlach, R. L. & Davis, L. E. *J. Vac. Sci Technol.* **14**, 339-342 (1977).
12. Loeb, G. I. & Baier, R. E. *J. Colloid. Interface Sci.* **27**, 38-45 (1968).
13. Blodgett, K. B. *J. Am. chem. Soc.* **57**, 1007–1022 (1935).
14. Mathewson, A. G. in *Proc. int. Symp. on Plasma-Wall Interactions* 517–536 (Pergamon, Oxford, 1977).
15. Powell, C. J. *Surface Sci.* **44**, 29–46 (1974).
16. Davis, L. E. MacDonald, N. C. Palmberg, P. W. Riach, G. E. & Weber, R. E. *Handbook of Auger Electron Spectroscopy* (Perkin-Elmer, Eden Prairie, 1976).
17. Dawson, P. H. *Int. J. Mass. Spectrom. Ion Phys.* **17**, 447–467 (1975).
18. Hagstrum, H. D. *Phys. Rev.* **119**, 940–952 (1960).
19. Weiss, R. *Rev. Scient. Instrum.* **32**, 397–401 (1961).

Comparison of sulphide deposits from the East Pacific Rise and Cyprus

E. Oudin, P. Picot & G. Pouit

Bureau des Recherches Geologiques et Minières, BP6009,
45018 Orléans Cedex, France

Submarine sulphide deposits of the East Pacific Rise, discovered in 1978, at the mouth of the Gulf of California are compared here with fossil massive sulphide ore bodies. Two main types of massive sulphide ore bodies are defined in relation to the associated volcanism. The Cyprus ore shows striking analogies with East Pacific Rise deposits. The comparison is based on mineralogical observations and textural facies evidence. These similarities suggest that the genetic processes were identical.

Cu–Zn–Pb–Ag–Au polymetallic massive sulphide deposits associated with submarine volcanics have formed throughout geological time: from the Archaean (North-west Canada) up to the Tertiary (Miocene Kuroko deposits of Japan). These deposits are often rich and may reach 10% (or more) in metal contents. Massive or layered ore is generally underlain by a stockwork of sulphides associated with a hydrothermal alteration pipe. Hydrothermal exhalative circulations on the ocean floor which are contemporaneous with the volcanism are, in fact,

Table 1 Mineralogical association

Deposit minerals	CYAMEX	ALVIN	Cyprus
Magnetite		T	
Haematite		T	
Pyrite	A	A to F	A
Marcasite	F	R	F to R
Pyrrhotite	T	R	T
Chalcopyrrhotite	F to R	F to R	
Chalcopyrite	F to R	F to R	F to R
Wurtzite or sphalerite	F	A	R
Bornite		F	R
Digenite	T	R	R
Chalcocite		T	R
Idaite		T	T
Covellite	T	T	T
Galena		T	T
Jordanite		T	T
Tennantite		T	T
Native Gold			T
Native Sulphur		R	
Goethite	A	T	T
Anhydrite	T	A	
Baryte	T	R	
Quartz			R
Opal	R	R	R
Carbonates		T	R

A, Abundant; F, frequent; R, rare; T, trace.

0028–0836/81/220404—04$01.00

the most characteristic common features of these deposits. Two types of massive sulphide deposits can be distinguished. The first shows a close association with calc-alkaline acid volcanics being part of acidic to basic differentiation[1]. The Archaean deposits in Canada, the Palaeozoic Huelva Province in Spain, and the Miocene Kuroko deposits of Japan are all of this type. The second type of copper rich mineralization exhibited by the Cyprus deposits, is associated only with basic tholeiitic volcanism and has less economic importance.

The two types have very different tectonic environments which explains the different nature of the volcanism involved. The first type is found in Island Arcs related to subduction zones (for example, the Kuroko deposits of Japan) and the second type is located in ophiolite complexes inferred to be the products of an accreting plate boundary (Mid-Ocean-Rise). Sedimentary features such as compositional banding are characteristic of the first type while collomorph textures[2] are typical of the second. The latter shows more important structural controls.

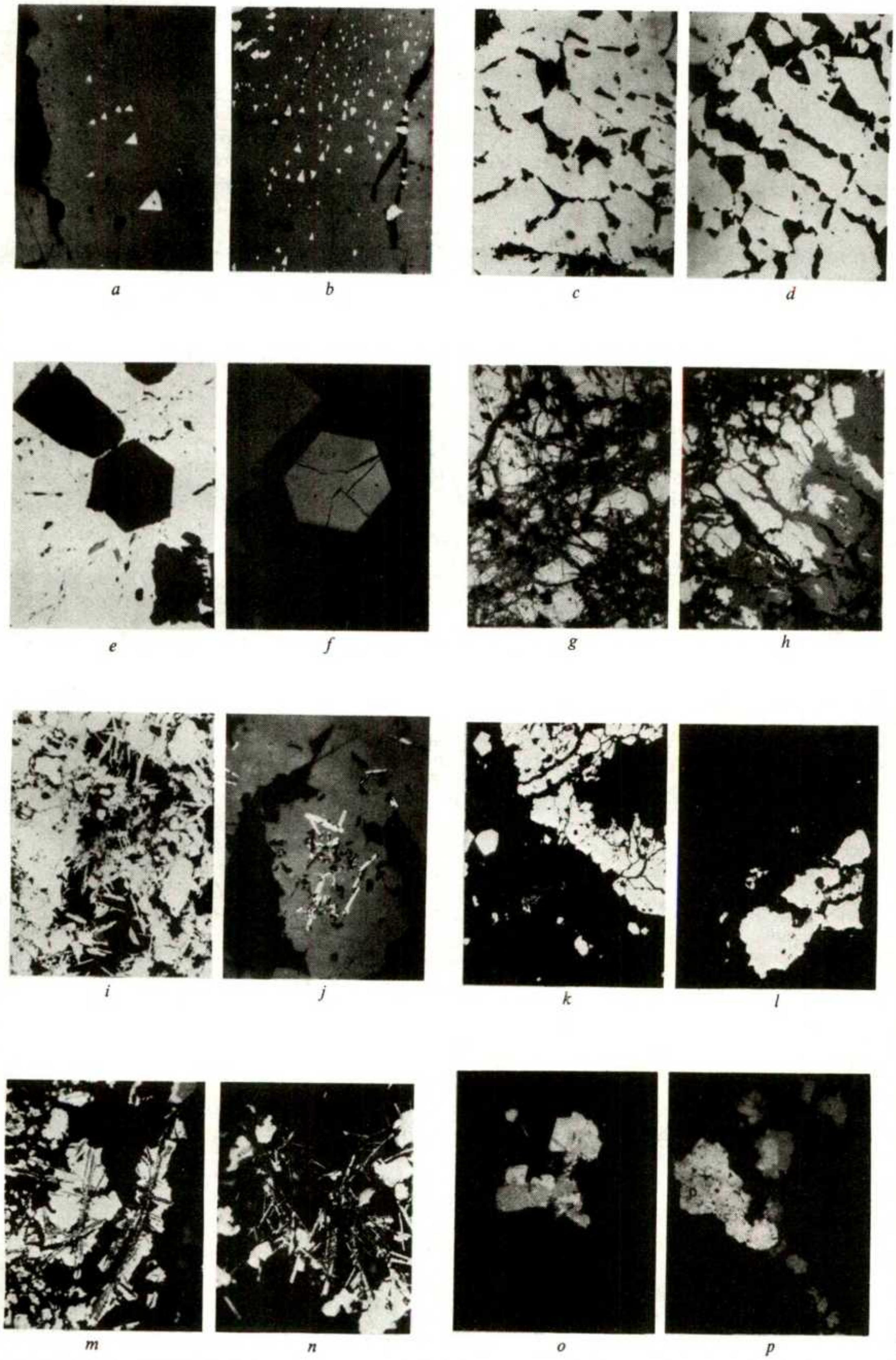


Fig. 1 *a*, triangular-shaped inclusions of chalcopyrite in wurtzite from Cyprus. Polished section, PPL × 200. *b*, Similar triangular-shaped inclusions of chalcopyrite in ALVIN wurtzite. Polished section, PPL × 160. *c*, Massive chalcopyrite formed by aggregates of euhedral grains. The interstitial spaces are open, from Cyprus. Polished section, PPL × 150. *d*, Similar massive chalcopyrite from ALVIN. Polished section, PPL × 150. *e*, Hexagonal crystal of ZnS (originally wurtzite) included in pyrite from Cyprus. Polished section, PPL × 500, in oil immersion. *f*, hexagonal wurtzite in ALVIN sample (included in araldite). Polished section, PPL × 160. *g*, Replacement of chalcopyrite by bornite, from Cyprus. Polished section, PPL × 150. *h*, Similar textures in ALVIN sample. Polished section, PPL × 150. *i*, Lamellar chalcopyrite on massive chalcopyrite (pyrite is whiter and harder) from Cyprus. Polished section, PPL × 150. *j*, Lamellar chalcopyrite included in the zinc sulphide (medium-grey) of ALVIN. Polished section, PPL × 600 oil immersion. *k*, Pyrite rimmed by opal, from Cyprus. Polished section, PPL × 150. *l*, Pyrite rimmed by opal, from ALVIN. Polished section, PPL × 160. *m*, Lamellar bornite embedded in pyrite, from Cyprus. Polished section, PPL × 150. *n*, Lamellar pyrrhotite partially altered and associated with some pyrite, from ALVIN. Polished section, PPL × 600 oil immersion. *o*, Jordanite (grey) associated with galena (white), from Cyprus. Polished section, PPL × 950 oil immersion. *p*, Jordanite (grey) surrounding galena (white), pyrite (Py), from ALVIN. Polished section, PPL × 950 oil immersion.

In Cyprus the mineralization occurs at the top of an ophiolitic series. It is situated in and above the lower pillow-lavas (LPL) underlying the upper pillow-lavas (UPL) which have the manganiferous umbers at the top. The pyritic ore forms a stockwork which becomes massive towards the top and is enriched in copper (1–10%) and locally in zinc. This massive mineralization is usually concentrated in basin-like depressions at the top of the LPL and is often coated by 'ochres' which are thought to be a submarine gossan and equivalent to submarine actual metalliferous sediments. Nearly all the deposits are at the intersection of important synsedimentary faults which have been reactivated after the ore deposition.

Relatively recent hydrothermal sulphide deposits were discovered in 1978 during the CYAMEX expedition in the mouth of the Gulf of California. They are situated at a depth of 2,600 m on the accreting plate boundary of the East Pacific Rise[6,7]. The sulphides occur in 'mounds' which are 10 m high and average 5 m in diameter. These chimneys are situated ~700–800 m away from the ridge axis and occur in small depressions (~30 m in width and depth). A later expedition with the American submersible ALVIN discovered active hydrothermal vents[8–10] creating massive sulphide mounds very close to the ridge axis. These samples are fresher and usually better crystallized than the CYAMEX ones. Sampling and their relative proximity to the rift axis might be partly responsible for the difference in textures and paragenesis. Abundant goethite and rare gypsum are characteristic of the CYAMEX samples which represent a submarine gossan.

Previous mineralogical studies of deep-sea sulphide deposits from the CYAMEX area[11,12] revealed that the samples were markedly altered. However, no comparison could be made with ore deposits associated with ophiolitic complexes such as Cyprus. New data from the study of unaltered samples from Alvin as well as new samples from Cyprus suggest that the genetic processes were similar for both recent (ALVIN)[13–16] and fossil deposits (Cyprus). This analogy can be substantiated from: (1) textural evidence and facies; and (2) similarities of mineral paragenesis and more specifically, the presence or absence of some trace minerals.

(1) Textural evidence and facies: neither deposit shows banding. In Cyprus the ore shows various facies: brecciated, disseminated, in stockwork and massive. In ALVIN the ore is massive and sometimes brecciated while in CYAMEX (altered). Collomorphic and arborescent (dendritic) crystalline growth can be observed in both cases although such textures are much more abundant in CYAMEX–ALVIN. In both, massive chalcopyrite is formed by the aggregation of euhedral grains. The interstitial spaces are open (Fig. 1*c*, *d*). Chalcopyrite also forms triangular inclusions in zinc sulphide which could represent an epitaxic growth on the wurtzite crystals, which are present in both deposits (Fig. 1*a*, *b*).

One can also observe: well developed hexagonal crystals of wurtzite in CYAMEX–ALVIN and probable sphalerite pseudomorphs after wurtzite in Cyprus ore (Fig. 1*e*, *f*). Wurtzite spherolites have now been identified in Cyprus ore. They are absent in ALVIN samples but common in CYAMEX where they crystallize *in situ* from hexagonal wurtzite. In both ALVIN and Cyprus, various mineral species (such as chalcopyrite, and bornite show lamellar structure which could represent a replacement of former platy minerals (Fig. 1*i*, *j*, *m*, *n*). Chalcopyrite is often replaced by bornite in both deposits (Fig. 1*g*, *h*). Bornite or pyrite are often coated by a thin rim of sphalerite. Pyrite, chalcopyrite or zinc sulphide are often coated by opal and sometimes rimmed by carbonates and are observed in Alvin and Cyprus deposits (Fig. 1*k*, *l*).

(2) Paragenetic similarities: Jordanite ($Pb_5As_2S_8$) which is usually found in very different settings (such as karst deposits, and deposits in the sedimentary cover) is a relatively rare mineral; however, it occurs in both mineral assemblages (Fig. 1*o*, *p*).

Trace minerals such as tellurides and cassiterite common in and characteristic of many volcanogenic[17] deposits are unknown in Cyprus and CYAMEX–ALVIN. However, these trace minerals seem to occur more often in Island Arc massive sulphide ore. Furthermore— the existence of finely expressed cassiterite seem to be related to the presence of acidic volcanism.

Finally some differences must also be considered. Chalcopyrrhotite[18,19] (a high temperature form of cubic cubanite) is completely absent in Cyprus. This can be understood if one considers that chalcopyrrhotite is an unstable phase which was preserved in CYAMEX–ALVIN samples as a result of quenching. Therefore chalcopyrrhotite is thought to be easily altered by diagenetic or low-grade metamorphic processes. Wurtzite, which is abundant in CYAMEX–ALVIN and rare in Cyprus, has a similar behaviour pattern probably reverting to sphalerite in fossil deposits. Pyrrhotite is abundant in ALVIN but scarce in CYAMEX where it changes to marcasite and iron sulphates. In Cyprus it is rare or altered to marcasite. As noted above, ALVIN is situated very near the rift axis whereas CYAMEX is 1 km away and can be considered as a fossil deposit in which the easily altered pyrrhotite has nearly all been transformed. Cyprus distance[6] to the fossil rift axis is 2.5 km.

The mechanism of formation of a hydrothermal ore deposit on the ocean floor has been described elsewhere[20–23] before the discovery of ALVIN hydrothermally active vents. Direct observation of the eruption of hot hydrothermal fluids on the ocean floor has provided important new data. From stable[24,25] and radiogenic isotope[26,27] measurements we know that seawater pervades the permeable oceanic basalts and circulation occurs as a result of the elevated temperatures associated with nearby magma chamber(s).

The efficiency with which hot water dissolves metals in basalts has been demonstrated elsewhere[28–32]. Partial chemical analyses[10,33] show that the hydrothermal fluids are enriched in calcium, lithium, silicium and manganese and depleted in magnesium. As predicted by Spooner[22], a localized sulphide precipitation occurs on the ocean floor as a result of a quenching phenomenon: when 350 °C (maximum) hydrothermal fluids erupts into the +2 °C ocean bottom water. The stockwork mineralization underlying the massive ore would represent the channel through which the hydrothermal fluids had ascended. Equivalent sulphide veins could be formed within the upper few hundred metres of the basaltic crust in ALVIN zone, as a result of adiabatic expansion of the hydrothermal fluids[33]. The ochre and umbers at the top of Cyprus ore bodies can be compared with recent metalliferous sediments found near submarine volcanic or rift zones in the Pacific[34].

The mineralogical assemblages and textures of ALVIN samples are directly related to their genetic process of formation. Low-grade metamorphic ore from Cyprus shows very striking analogies which suggest a similar genetic process. The differences in mineralogy are readily understandable when one considers the alteration and diagenetic processes. These results confirm the distinctions based on tectonic environment and ore mineralogy between the two types of massive sulphide deposits. (1) Mineralizations associated with calc-alkaline volcanics, where trace minerals such as cassiterite and tellurides (mainly hessite and altaite) are frequent. (2) Mineralizations associated with basic tholeiitic volcanism where cassiterite is never observed and tellurides are absent, as in Cyprus and CYAMEX–ALVIN.

However, deposits with frequent tellurides[35] are present in ophiolite series but are restricted to the basic to ultrabasic lower unit[36]. The tellurides observed in such deposits are hessite and melonite (a nickel telluride which indicates a basic influence). Lead telluride (altaite) is absent and clausthalite (a lead selenide) is formed instead.

We thank G. Constantinou and R. Hekinian for collaboration and assistance.

Received 10 October 1980; accepted 23 March 1981.

1. Pouit, G. *Mém. Soc. geol. Fr.* 115–126 (1976).
2. Adamides, N. G. *Trans. Instn Min. Metal.* **84,** B17–B23, 317–322 (1975).
3. Constantinou, G. & Covett, G. J. S. *Trans. Instn Min. Metal.* **81,** B34–B46, 334–345 (1971).
4. Constantinou, G. *Econ. Geol.* **68,** 83–86 (1973).

5. Hutchinson, R. W. & Searle, D. F. *Soc. Miner. Geol. Jap. Spec. Iss.* **3**, 198–205 (1971).
6. CYAMEX Scientific team *EOS* **59**, 1198 (1978); *C.r. hebd. Acad. Séanc. Sci., Paris* 287 (1978).
7. Francheteau, J. *et al. Nature* **277**, 523–528 (1979).
8. MacDonald, K. C. *et al. Earth planet. Sci. Lett.* **48**, 17 (1980).
9. Edmond, J. M. *et al. EOS* **60**, 864 (1979).
10. Spiess, F. N. *et al. Science* **207**, 1421–1433 (1980).
11. Hekinian, *et al. Science* **207**, 1433–1444 (1980).
12. Picot, P. *et al. Docum. Bur. Rech. Geol. Min. No.* 20 (1980).
13. Oudin, E. *Docum. Bur. Rech. Geol. Min.* (in the press).
14. Haymon, R. *et al. EOS* **60**, 864 (1979).
15. Haymon, R. M. & Kastner, M. *Descriptive Catalog of Rise Sulfide Samples from the East Pacific Rise, 21 °N*, 510 (Reference Series 79–28, 1979).
16. Haymon, R. *et al. Earth planet. Sci. Lett.* (in the press).
17. Aye, F. *et al. Bull. Miner.* **101**, 139–147 (1978).
18. Cabri, J. L. *et al. Can. Miner.* **12**, 33–38 (1973).
19. Cabri, J. L. *Econ. Geol.* **68**, 443–454 (1973).
20. Ohmoto, H. *et al. Econ. Geol.* **69**, 947–953 (1974).
21. Bonatti, E. A. *Rev. Earth planet. Sci. Lett.* **3**, 401–431 (1975).
22. Sponner, E. T. C. *et al. Contr. Miner. Petrol.* **42**, 287–304 (1973).
23. Spooner, E. T. C. *Trans. Instn Min. Metal.* 58–68 (1976).
24. Heaton, T. H. E. *et al. Trans. Instn Min. Metal.* 42–57 (1976).
25. Sheppard, S. M. F. *Trans. Instn Min. Metal.* 25–41 (1976).
26. Spooner, E. T. C. *et al. Geochim. cosmochim Acta* **41**, 873–890 (1977).
27. Chapman, H. J. *et al. Earth planet. Sci. Lett.* **35**, 71–78 (1978).
28. Bischoff, J. L. *et al. Earth planet. Sci. Lett.* **25**, 385–397 (1975).
29. Hajash, A. *Contr. Miner. Petrol.* **53**, 205–226 (1975).
30. Bischoff, J. L. *et al. Proc. 2nd int. Water–Rock Interaction Symp.* Strasburg (1977).
31. Seyfried, W. *et al. Earth planet. Sci. Lett.* **34**, 71–77 (1977).
32. Mottl, *et al. Geochim. cosmochim. Acta* **43**, 869–884 (1978).
33. Bischoff, J. L. *Science* **207**, 1465–1469 (1980).
34. Guillemot, D. *et al. Proc. Int. Ophiolite Symp.*, Cyprus (1980).
35. Bouladon, J. *et al. Bull. Bur. Rech. Geol. Min.* No. 1, 24–41 (1968).
36. Constantinou, G. *Proc. int. Ophiolite Symp.*, Cyprus (1980).

Are high rates of sulphate reduction associated with anaerobic oxidation of methane?

Allan H. Devol

Fisheries Research Institute, WH-10, University of Washington, Seattle, Washington 98195, USA

Saiyed I. Ahmed

Department of Oceanography, WB-10, University of Washington, Seattle, Washington 98195, USA

Classical models of sulphur diagenesis in marine sediments are based on the assumption that the rate of sulphate reduction is first order with respect to oxidizable particulate organic carbon (POC). This assumption requires that oxidizable POC, sulphate concentration and the sulphate reduction rate be highest at the top of the sulphate reduction zone and decrease exponentially with increasing sediment depth[1,2]. However, to explain recent observations of concave upwards methane distributions, the anaerobic consumption of methane has been proposed. Furthermore, it has been proposed that this consumption takes place near the bottom of the sulphate reducing zone where sulphate concentrations are low[3–6]. Thus, if sulphate reducing bacteria are associated with the anaerobic oxidation of methane, a peak in sulphate reduction rate might be expected in this deep consumption zone[7]. The importance of the process in sedimentary sulphur diagenesis is indicated by calculations estimating that 30–75% of the downward sulphate flux at depth may be consumed by methane oxidation within this zone[6,8,9]. We present here profiles of sulphate reduction rate in anoxic sediments that show distinct local maxima at the depth where the anaerobic oxidation of methane would be expected. Our measurements were made during July and August 1978 in Saanich Inlet, an anoxic fjord located on the south-east side of Vancouver Island, British Columbia. The inlet has a shallow sill (~70 m) which restricts circulation of the deeper water (maximum depth 225 m) inside the basin to the extent that for about 8 months of the year the bottom waters contain no oxygen, nitrate or nitrite, and significant concentrations of free H_2S are present[10]. As the deep waters contain hydrogen sulphide, the inlet is an ideal location for studying sedimentary sulphate reduction because reactions with oxygen and the effects of burrowing organisms can be neglected.

All sediment samples were taken at a station located in 225 m of water using a hydraulically damped multiple corer that avoids disturbance of the sediment water interface[11]. Typical distributions of SO_4^{2-} and total dissolved sulphide reflect the strongly reducing sedimentary environment in Saanich Inlet (Fig. 1*a*). Sulphate was exhausted at depths of ~20–30 cm in the sediment column and sulphide concentrations increased rapidly with depth, reaching maximum concentrations near the depth of sulphate depletion. Below this depth dissolved sulphide concentrations decreased. This type of dissolved sulphide profile is characteristic of strongly reducing sediments and is due to sulphide production in the zone where sulphate is present and removal as FeS below[12]. Sediment porosity was high at the surface, ~0.96 cm^3 per cm^3 and decreases only slightly in the upper 40 cm of the sediment column (Fig. 1*b*). At any given depth sulphate reduction rates were somewhat variable: however, the main features of the profiles were consistent from core to core. In general, the highest rates were observed at the surface and decreased with increasing depth until ~10 cm. Below this depth a local maximum was observed in all but one core (Fig. 2).

The profiles presented here cannot be modelled within the framework of current diagenetic models. Assuming steady state and no absorption of sulphate by sediment particles, the distribution of sulphate or any other soluble species can be described by equation (1)[1]:

$$\frac{\partial}{\partial z}\left(\phi D\frac{\partial[\mathrm{SO_4}]}{\partial z}\right)-\frac{\partial}{\partial z}(\phi\omega[\mathrm{SO_4}])+R=\frac{\partial[\mathrm{SO_4}]}{\partial t}=0 \qquad (1)$$

where ϕ is the porosity, ω is the sedimentation rate, z is depth taken positively downward, t is time, D is the apparent diffusion coefficient of the sulphate in sediment, and R is the sulphate reduction rate. It is generally assumed that the sulphate reduction rate is first order with respect to oxidizable POC. Thus, because POC is nondiffusible, we can also write

$$-\omega\frac{\partial G}{\partial z}=kG=-\frac{R}{\alpha} \qquad (2)$$

where G is the concentration of metabolizable POC, α is a stoichiometric coefficient equal to the number of sulphate ions reduced per atom of carbon oxidized and k is a first-order rate constant[1]. Assuming constant values of ω, α and k, the solution to equation (2) with the boundary conditions that $G = G_0$ at

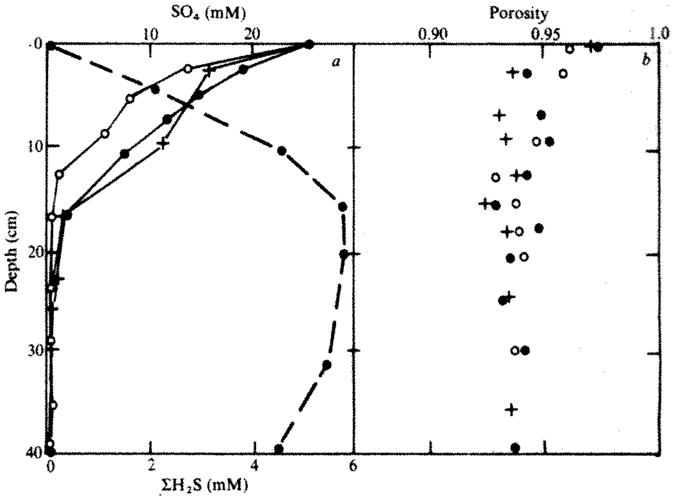


Fig. 1 Typical profiles of sulphate (solid lines, *a*), total dissolved sulphide (dashed line, *a*) and porosity (*b*). For sulphate and total dissolved sulphide, cores were sectioned and the pore water expressed into nitrogen filled test tubes using a gas (N_2) operated squeezer. Dissolved sulphide analyses[14] were performed immediately, while sulphate samples were frozen for later analyses[15]. Bulk sediment porosity was determined by weight loss after drying at 80 °C and corrected for salt content. Symbols distinguish measurements made on replicate cores.

0028-0836/81/220407—02$01.00

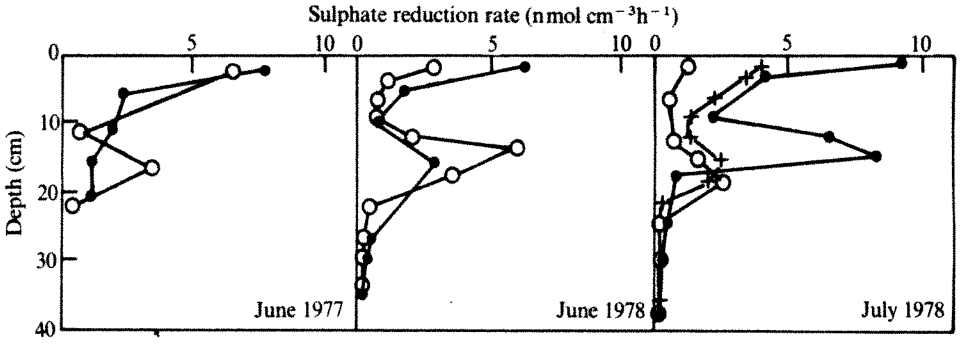


Fig. 2 Profiles of sulphate reduction rate. Sulphate reduction rates were determined on whole sediments using the method of Jørgensen[16] as modified by Nedwell and Abram[17]. Symbols distinguish measurements made on replicate cores.

$z = 0$ and G approaches zero as z approaches infinity is $G = G_0 \exp[-(k/\omega)z]$. This solution requires that the concentration of metabolizable POC and therefore the sulphate reduction (αkG) decrease exponentially with sediment depth. This is clearly not the case for the sulphate reduction rate data from Saanich Inlet.

That the sulphate reduction rate does not decrease exponentially as predicted by the model indicates that there is either a corresponding increase in oxidizable POC at ~15 cm, or that other, presumably dissolved, substrates are being utilized. Although there is no method available to determine metabolizable POC, total POC has been determined at our sampling station[9] and no subsurface peak is evident. In fact, the total POC data can be fitted reasonably well with an exponentially decreasing oxidation rate function (Fig. 3*b*).

To investigate the possibility that dissolved organic carbon (DOC) was the substrate responsible for the subbottom peak in the sulphate reduction rate, we also measured the DOC concentration at various depths within the sediments (Fig. 3*a*). However, because of the analytical method used, our DOC measurements do not include volatile organic components such as CH_4 (ref. 13). DOC concentration increased steadily with depth and showed no apparent consumption in the region of the sulphate reduction peak. The DOC distribution is best modelled if we assume an exponentially decreasing net production rate.

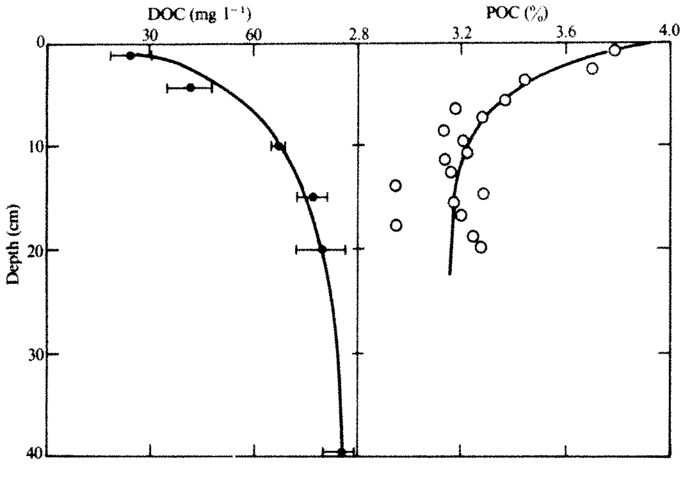


Fig. 3 Dissolved organic carbon (DOC) and total particulate organic carbon (POC) in Saanich Inlet sediments. DOC was determined by wet oxidation[13]. Samples were acidified, and bubbled with nitrogen to remove dissolved inorganic carbon. Consequently, volatile components such as CH_4 were also removed. The total POC profile is redrawn from ref. 9[9]. Both the DOC and total POC profiles were modelled using equation (1). For the general case we substitute C for $[SO_4]$ in equation (1), where C represents either DOC or total POC. The best fit to both the DOC and total POC data was obtained by assuming an exponential rate term. ($P = P_0 \exp[-\beta z]$, where β is a decay coefficient and P is substituted for R in equation (1). For DOC, P represents a net production rate which decreases exponentially from a surface value of P_0. For POC, P represents an oxidation which decreases exponentially from a surface value of P_0). Given the boundary conditions, $C = C_0$ at $z = 0$ and $C \rightarrow C_\infty$ as $z \rightarrow \infty$ (where C_∞ is finite), and assuming the values of ω and D presented by Murray *et al.*[9], the solution to the general form of equation (1) with an exponential rate term is $C = (C_0 - C_\infty)\exp[-\beta z] + C_\infty$, where $(C_0 - C_\infty) = P_0/(D\beta^2 + \omega\beta)$ (ref. 15). The model fits are shown as solid lines.

Thus, both the total POC and nonvolatile DOC profiles can be modelled as smooth functions of depth, but, within the context of classical diagenetic models it is impossible to explain the subsurface peak in sulphate reduction rate because the models predict an exponentially decreasing rate function. However, the peak in the sulphate reduction rate profile is located in the depth zone where sulphate concentrations approach zero. This is exactly the depth zone where the anaerobic oxidation of methane, which would not be included in our DOC measurement, has been proposed to take place. Investigations in other marine sediments have shown that a significant fraction of the downward sulphate flux at depth is consumed within the anaerobic methane oxidation zone[6,8]. That the magnitude of anaerobic methane oxidation is sufficient to support the sulphate reduction peak observed in Saanich Inlet is indicated by a previous study in which an areal methane oxidation rate of 7.7×10^{-12} mol cm^{-2} s^{-1} was calculated[9]. Assuming oxidation takes place through the reaction, $CH_4 + SO_4^{-2} \rightarrow HS^- + HCO_3^- + H_2O$, and that the reaction occurs principally in a 10 cm thick depth zone where sulphate concentrations approach zero, this areal methane oxidation rate would be equivalent to a volumetric sulphate reduction rate of 2.8 nmol cm^{-3} h^{-1}. This value is in good agreement with the observed data (Fig. 2). We therefore conclude that the high rates of sulphate reduction we observe at depth are associated with the anaerobic oxidation of methane, either directly or indirectly, through reactions coupled to other substrates.

We thank J. W. Murray and J. P. Christensen for critically reading the manuscript. This research was supported by NSF Grants OCE 78-10773 and OCE 79-19176. Contribution 554 from the College of Fisheries, University of Washington.

Received 22 December 1980; accepted 23 March 1981.

1. Berner, R. A. in *The Sea*, Vol. 5 (ed. Goldberg, E. D.) 427–450 (Wiley-Interscience, New York, 1974); *Early Diagenesis* (Princeton University Press, New Jersey, 1980).
2. Lerman, A. in *The Sea* Vol. 6 (ed. Goldberg, E. D.) 695–738 (Wiley-Interscience, New York, 1977).
3. Barnes, R. D. & Goldberg, E. D. *Geology* **4**, 297–300 (1976).
4. Reeburgh, W. S. & Heggie, D. T. *Limnol. Oceanogr.* **22**, 1–9 (1977).
5. Martins, C. S. & Berner, R. A. *Limnol. Oceanogr.* **22**, 10–25 (1977).
6. Reeburgh, W. S. *Earth planet. Sci. Lett.* **47**, 345–352 (1980).
7. Mountfort, D. O., Asher, R. A., Mays, E. R. & Tiedje, J. M. *Appl. Envir. Microbiol.* **39**, 686–694 (1980).
8. Reeburg, W. S. in *The Dynamic Environment of the Ocean Floor* (eds Fanning, K. A. & Manheim, F. T.) (Heath, Boston, in the press).
9. Murray, J. W., Grundmanis, V. & Smethie, W. M. Jr. *Geochim. cosmochim. Acta* **42**, 1101–1026 (1976).
10. Anderson, J. J. & Devol, A. H. *Estaur. Coast. mar. Sci.* **1**, 1–10 (1973).
11. Pamatmat, M. M. *Limnol. Oceanogr.* **16**, 536–550 (1971).
12. Goldhaber, M. B. & Kaplin, I. R. *Soil Sci.* **119**, 42–55 (1975).
13. Menzel, D. W. & Vaccaro, R. F. *Limnol. Oceanogr.* **9**, 139–142 (1964).
14. Cline, J. D. *Limnol. Oceanogr.* **14**, 454–458 (1969).
15. Jørgensen, B. B. *Geomicrobiology J.* **1**, 11–27 (1978); 29–47 (1978).
16. Macchi, G. R., Cescon, B. S. & Mameli-D'Errico, D. *Arch. Oceanogr. Limnol.* **16**, 163–171 (1969).
17. Nedwell, D. B. & Abram, J. W. *Estaur. Coast. mar. Sci.* **6**, 341–351 (1978).

Albumin systematics of the extinct mammoth and Tasmanian wolf

Jerold M. Lowenstein*, Vincent M. Sarich† & Barry J. Richardson‡

* Department of Medicine (122 MR2), University of California, San Francisco, California 94143, USA
† Departments of Anthropology and Biochemistry, University of California, Berkeley, California 94720, USA
‡ Department of Population Biology, The Australian National University, Canberra City, ACT 2600, Australia

Knowledge of the molecular systematics of living species[1-3] has provided a framework, independent of morphology, for evaluating the genetic relationships of living forms. Although amino acids have been found in many animal and plant fossils[4], genetic information has generally not been obtained from the small amounts of surviving, chemically degraded protein. However, Westbroek *et al.*[5] have described immunological reactions to material from 60-Myr-old molluscs, and Lowenstein[6,7] has reported the identification by radioimmunoassay (RIA) of species-specific collagen and serum factors in primate and bovine fossils. We report here the use of RIA to detect and characterize albumin in the soft tissues of two recently extinct species, the Siberian mammoth (*Mammuthus primigenius*) and the Tasmanian wolf (*Thylacinus cynocephalus*). Mammoth albumin was found to be very similar to, and immunologically equidistant from, the albumins of the two living species of elephants, Indian (*Elephas maximus*) and African (*Loxodonta africana*). Tasmanian wolf albumin stands in the same relationship to those of the two Australian marsupial dasyurids, *Dasyurus* and *Dasyuroides*.

Our RIA is a double-antibody radioimmunometric assay with polyvinyl microtitre plates as the solid-phase ligand[6,8]. Fossil extracts, or protein standards, are left for 1 h in the microtitre cups, where some protein binds irreversibly to the plastic. The rest is washed out with a solution containing 2% chicken egg protein to block any further binding of plastic to protein. Next, rabbit antisera against albumins[9] are added to the cups and left at room temperature for 24 h; some antibody binds to the plastic-bound albumin, the rest is washed out. About 1 ng ^{125}I-labelled goat anti-rabbit γ-globulin (GARGG) is then added to each cup and left for 24 h. The amount of GARGG bound increases with the amount of rabbit antibody present. The unbound radioactive solution is washed out, the plate cut up and individual cups assayed for radioactivity in a scintillation counter. Final radioactivity is a function of both the amount of albumin on the cup and the specificity of the first antiserum for that albumin. Sensitivity of the method (smallest concentration of albumin detectable) is ~1 ng ml^{-1}—about one-thousandth the amount detectable with the sensitive microcomplement fixation (MC′F) technique which has been used extensively for studying albumin evolution[3,9-12]. In the 'direct' RIA described here, specificity (discrimination among closely related species) is, however, inferior to that of MC′F or quantitative precipitin analysis.

Table 1 Direct RIA reactions of mammoth and other albumins

Antigen	% Uptake of radioactive second antibody			
	IE	AE	Mn	H
Mammoth (*Mammuthus primigenius*)	10.0	10.3	0.25	0.20
Indian elephant (IE) (*Elephas maximus*)	10.0	10.3	0.84	3.2
African elephant (AE) (*Loxodonta africana*)	9.7	10.0	0.83	3.1
Manatee (Mn) (*Trichechus manatus*)	2.1	2.2	10.0	2.8
Human (H) (*Homo sapiens*)	2.1	2.2	0.80	10.0

With rabbit anti-albumin concentration fixed at 10^{-4} M, homologous albumin concentration in the plastic cup was adjusted to give 10% uptake of labelled GARGG. Mammoth extract was adjusted to give 10% uptake with anti-IE albumin. The mammoth extract reacts far more specifically with antisera to elephant albumin than to the others. Reciprocal values are not particularly good with this method, and the cross-reactions fail to show the relationship (albeit distant) of the manatee with the elephants. The two elephant albumins react identically.

Table 2 Immunological similarities and distances among mammoth, elephant and other albumins

Antigen	Inhibition IE	AE	Mn	B	H
Mammoth	0.94	0.92	0.30	0.05	0.03
Indian elephant (IE)	1.0	0.92	0.33	0.10	0.07
African elephant (AE)	0.94	1.0	0.45	0.10	0.07
Manatee (Mn)	0.39	0.40	1.0	0.06	0.12
Bison (B)	0.33	0.33	0.20	1.0	0.12
Human (H)	0.22	0.22	0.23	0.01	1.0

For all albumins but mammoth, inhibition was achieved by premixing inhibitor serum (10^{-2} M) with antiserum (10^{-4} M) homologous to the solid-phase albumin (10^{-5} M). Thus inhibitor exceeded bound antigen by a factor of 10^3. 1.0 indicates complete inhibition (no uptake of radioactive antibody) and immunological identity; 0 would represent no inhibition and no immunological similarity. For the mammoth, where the amount of albumin available was insufficient for use as an inhibitor and clean antisera to it were not available[13], the values were obtained by comparing the ability of its plastic-bound albumin to bind the various antisera, inhibited and uninhibited, with that shown by the other albumins. For example, AE albumin inhibits the reaction between mammoth albumin and anti-IE albumin no more or less strongly than that between that antiserum and IE albumin. Thus mammoth albumin is as distant from that of AE as is IE albumin. The reciprocal reaction, using IE albumin as the inhibitor for anti-AE, tells you that mammoth albumin is also as distant from IE albumin as is that of AE. The same logic follows for all the other comparisons.Reciprocal values are fairly good, decreasing in reliability as immunological similarity decreases. The two elephant species are clearly differentiated in this approach, so that we were able to identify seven out of seven different 'unknown' sera, four Indian and three African. The relation of the manatee to the elephants is also apparent.

RIA specificity can be greatly improved by having an albumin in solution in the cup competing with the solid-phase albumin for available antibody. This then becomes an inhibition or competitive binding technique[8]. Soluble rabbit anti-albumin is premixed with excess albumin or serum (100–1,000 times the concentration originally applied to the plastic cup), and binds essentially all the antibody directed against common determinants of the two competing albumins, permitting only those antibody populations which react more strongly with the determinants on the plastic-bound albumins to bind to the latter. The more similar the two competing albumins, the less will be the uptake of antibody (and hence radioactivity) on the plastic.

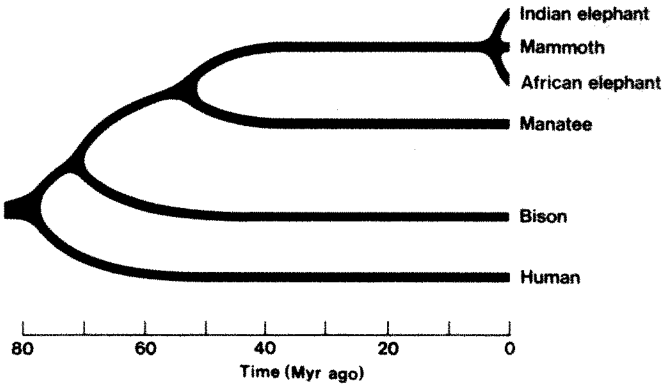


Fig. 1 An albumin phylogeny of the mammoth and related forms, based on the data of Tables 1 and 2 and ref. 13. Those data suggest relatively similar amounts of albumin evolution along the various lineages represented and thus we feel justified in adding the approximate time scale included. We assume a divergence time of 80 Myr between the ungulate and non-ungulate placentals.

0028–0836/81/220409—03$01.00

Table 3 Direct RIA reactions of Tasmanian wolf and other marsupial albumins

Antigen	% Uptake of radioactive second antibody D	Dd	My	E	Be
Tasmanian wolf (*Thylacinus*)	9.1	10.0	8.3	6.8	4.0
Dasyurus (D)	10.0	8.5	8.3	4.8	3.5
Dasyuroides (Dd)	10.9	10.0	8.9	6.0	4.3
Myrmecobius (My)	7.7	5.2	10.0	3.7	2.0
Echymipera (E)	5.8	5.5	4.9	10.0	3.8
Bettongia (Be)	3.5	2.6	3.4	4.9	10.0
Caluromys	2.5	1.1	1.5	1.1	0.4
Human	2.1	0.6	0.9	0.7	0.3

The direct reactions of Tasmanian wolf muscle extract with various marsupial anti-albumins show that it groups with the dasyurids, and among those is slightly closer to *Dasyurus* and *Dasyuroides* than to *Myrmecobius*. In order of increasing distance are the peramelid *Echymipera* and the macropodid *Bettongia*, both Australian genera. Further still are the South American woolly opossum *Caluromys* and the placental mammals represented by human serum. As described in Table 1 legend, antiserum concentration was fixed at 10^{-4} M and serum concentration adjusted to give homologous uptake of 10%. Even with the direct reaction, reciprocity is good enough so that branching orders are evident from inspection.

We used this method to distinguish the albumins of the African and Indian elephant, which are indistinguishable by MC′F, and to characterize the similarities of the fossil albumins with albumin from related species. Each determination of antibody binding was run in duplicate or triplicate, with about 5% variation from run to run. Each study of the fossil albumins was run several times, with essentially identical results. Although only a small proportion of the GARGG added may be bound, especially in inhibited reactions, the amount remains statistically above that of the 'blank', consisting of cups coated with chicken egg protein. With the direct RIA method, maximum binding of antibody is generally achieved with antiserum concentration about 10^{-3} M; but maximal inhibition occurs at lower concentrations, usually 10^{-4} M. To calibrate the RIA results, we compared them with a standard method for determining evolutionary distances— MC′F—which has been tested on hundreds of albumins. With various albumin immunological distances, we found a consistent correlation between the results of the two methods; we therefore felt confident in applying this calibration to the fossil proteins.

The 40,000-yr-old baby Siberian mammoth known as Dima was discovered near Magadan, USSR in 1977. A specimen of right thigh muscle was obtained, ground in buffer and tested for albumin by MC′F, immunodiffusion and RIA. A portion was also injected into rabbits to make antisera[13]. Initially, neither MC′F nor immunodiffusion detected mammoth albumin directly, but after its presence had been shown by RIA [at ~1 μ ml^{-1} (or ~1% the amount extractable from fresh muscle)], it was also demonstrated, after further concentration, by the other methods, and was shown to be very similar to elephant albumin[13]. Antisera raised against the mammoth muscle suspension gave the strongest reactions with elephant albumin[13].

Direct RIA mammoth reactions are given in Table 1. Neither direct RIA, MC′F, nor immunodiffusion distinguishes between the albumins of the Indian and the African elephant. To determine whether mammoth albumin is immunologically closer to one or the other of these, we first tested seven different elephant sera as 'unknowns' and identified correctly the four Indian and three African sera. Next we compared the mammoth to these two species and found its albumin immunologically equidistant from them, giving a 93% cross-reaction with each using the inhibition technique described above (Table 2). The inhibition procedure apparently exaggerates the small difference between the two elephant albumins; judging by MC′F; or quantitative precipitin analysis they are >99% identical. The next most similar albumin, that of the manatee (*Trichechus*), gives only a 40% cross-reaction, while *Bison* and *Homo* albumins are at 22% and 15%, respectively. These data and their MC′F homologues facilitate the construction of an albumin phylogeny (Fig. 1) which corresponds well to current ideas about the relationships of these species[14]. The molecular distance between the two elephants when assessed by RIA inhibition, MC′F comparisons of albumin and transferrins, and electrophoretic comparisons of the serum proteins, is, however, only about half that for the corresponding proteins in humans, chimpanzees and gorillas. The divergence time of the latter trio, on the biochemical evidence, is ~5 Myr (refs 10, 11) which implies a shorter divergence time for the elephants, contrary to prevailing opinion[14].

While no one, to our knowledge, has suggested that the mammoth is not closely related to living elephants, the affinity of the recently extinct Tasmanian wolf with the dasyurids, the Australian marsupicarnivores with which it is generally grouped, has been seriously questioned. The marked morphological convergence of this species on dogs and foxes, its corresponding departure from any other Australian form, and strong cranial, dental and skeletal resemblances to the extinct South American borhyaenids have suggested[15–17] affinities to that group rather than the dasyurids.

We obtained three museum specimens of thylacine material: dried muscle (0.4 g) adhering to a bone collected before 1893 in Tasmania and two untanned skin samples of similar weight collected early in the twentieth century. Each specimen was finely ground in a 0.2 M EDTA buffer, *p*H 7.4, and extracted for 48 h. Extracts were tested by RIA with antisera to the albumins of *Dasyurus* (tiger cat), *Dasyuroides* (kowari), *Myrmecobius*

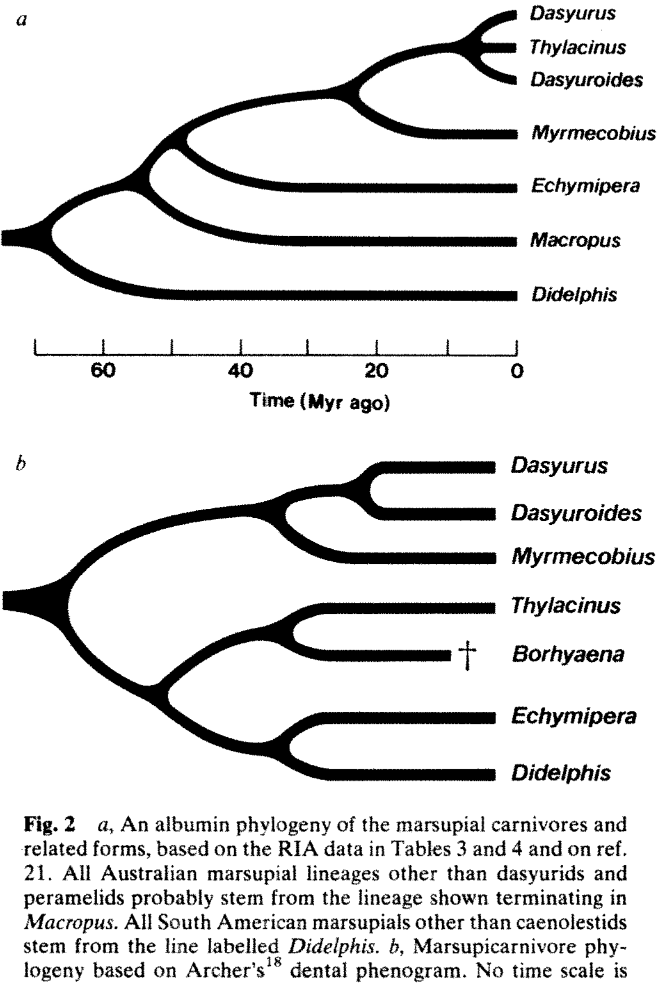


Fig. 2 *a*, An albumin phylogeny of the marsupial carnivores and related forms, based on the RIA data in Tables 3 and 4 and on ref. 21. All Australian marsupial lineages other than dasyurids and peramelids probably stem from the lineage shown terminating in *Macropus*. All South American marsupials other than caenolestids stem from the line labelled *Didelphis*. *b*, Marsupicarnivore phylogeny based on Archer's[18] dental phenogram. No time scale is implied. Note the markedly different placements of the *Thylacinus* and *Echymipera* lineages in this scheme, compared with that in *a*. *Borhyaena* is a long-extinct form not yet subject to biochemical study.

Table 4 Immunological similarities and distances of Tasmanian wolf and related marsupials

	Inhibition T	D	Dd	My	E
Thylacinus (T)		0.86	0.82	0.67	0.30
Dasyurus (D)	7		0.84	0.59	0.28
Dasyuroides (Dd)	9	8 (11)		0.71	0.32
Myrmecobius (My)	17	23 (41)	15 (39)		0.29
Echymipera (E)	52	55 (94)	49 (76)	54 (89)	

Averages of reciprocal inhibition similarities ($\mathrm{RIA}^{\mathrm{IS}}$) are given above the diagonal, except for *Thylacinus*, for which no anti-albumin was available. As with the mammoth, its similarities were obtained by inhibiting antisera to the albumins of the most closely related species, *Dasyurus* and *Dasyuroides*. Anti-D was inhibited with Dd serum and vice versa, giving 86% and 82% inhibition, respectively. The $\mathrm{RIA}^{\mathrm{ID}}$ (inhibition difference) values (below the diagonal) are calculated from the $\mathrm{RIA}^{\mathrm{IS}}$ values (above the diagonal) according to the equation $\mathrm{ID} = -100 \log_{10} \mathrm{IS}$. The corresponding MC′F values are given in parentheses. Clearly, the $\mathrm{RIA}^{\mathrm{ID}}$ values here are slightly more than half the corresponding MC′F values and a wider range of comparisons suggests that $\mathrm{MC'F}^{\mathrm{ID}} = 1.7\ \mathrm{RIA}^{\mathrm{ID}}$. Thus, the $\mathrm{RIA}^{\mathrm{ID}}$ values approximate the divergence times in Myr between the pairs of species involved.

(numbat), *Echymipera* (bandicoot) and (*bettongia* (rat kangaroo or bettong). The first three are dasyurids, the fourth a peramelid and the last a macropodid. The sera had been prepared for a project previously reported[21]. The three thylacine extracts gave identical species-specific cross-reactions, reacting as though they contained ~1% of the albumin that would be found in extracts of fresh tissues. This result is very similar to that obtained from the mammoth muscle and suggests that a small fraction of the albumin in soft tissues is sequestered and protected from leaching and denaturants. After extraction, the insoluble muscle tissue could also bind antibodies to marsupial albumins, indicating the presence of immunologically active albumin in this residue. We were able to extract 10 times more albumin from muscle than from skin and used this material for detailed comparisons.

Antisera to *Dasyurus* and *Dasyuroides* albumins reacted maximally with the thylacine extract (Table 3), indicating close immunological similarity. Progressively weaker reactions were seen with antisera to *Myrmecobius, Echymipera* and *Bettongia* albumins. These relationships were confirmed and quantified by inhibition reactions (Table 4), which showed *Dasyurus, Dasyuroides* and *Thylacinus* to be immunologically equidistant, the others more distant. The other largish dasyurid, the Tasmanian devil (*Sarcophilus*), gives a MC′F albumin distance to *Dasyurus* of only six units, about half that given by *Dasyuroides* and, by implication, *Thylacinus*. We therefore suggest that the Tasmanian wolf is part of a quite recent dasyurine radiation and separated from the lineages leading to *Dasyurus–Sarcophilus* and *Dasyuroides* in the late Miocene some 6–10 Myr ago (Fig. 2). The positions of *Thylacinus* and *Echymipera* differ profoundly in our scheme (Fig. 2*a*) and that of Archer (Fig. 2*b*).

These results should help settle the controversy about thylacine affinities. We agree with Simpson[19] and Marshall[20] that thylacine resemblances to borhyaenids must be due to parallel or convergent evolution, for the thylacine albumin is that of a very recently derived dasyurid. This may be the first of many phylogenetic issues to be affected by the protein immunology of extinct forms.

We thank the L.S.B. Leakey Foundation and the Wenner-Gren Foundation for Anthropological Research for support (to J.M.L.), Gary Scheuenstuhl for technical assistance, and Allan Wilson for mammoth tissue and laboratory facilities. Elephant sera came from the San Diego and Los Angeles zoos, marsupial sera were collected during an *Alpha Helix* expedition to New Guinea and from the Taronga Park Zoo. Tasmanian wolf samples were obtained through the cooperation of curators at the Australian Museum, National Museum of Victoria and Queen Victoria Museum in Launceston.

Received 14 August 1980; accepted 20 March 1981.

1. Ayala, F. J. (ed.) *Molecular Evolution* (Sinauer, Sunderland, Massachusetts, 1976).
2. Dayhoff, M. O. *Atlas of Protein Sequence and Structure* Vol. 5, Suppl. 3, (National Biomedical Research Foundation, Washington DC, 1978).
3. Wilson, A. C., Carlson, S. S. & White, T. J. *A. Rev. Biochem.* **46,** 573–639.
4. Wyckoff, R. W. G. *The Biochemistry of Animal Fossils* (Scientechnica, Bristol, 1972).
5. Westbroek, P. H. *et al. Paleobiology* **5,** 151–167 (1979).
6. Lowenstein, J. M. in *Biogeochemistry of Amino Acids* (ed. Hare, P. E.) 41–52 (Wiley, New York, 1980).
7. Lowenstein, J. M. *Naturwissenschaften* **67,** 343–346 (1980).
8. Tsu, T. T. & Herzenberg, L. A. in *Selected Methods in Cellular Immunology* (eds Mishell, B. B. & Shiigi, S. M.) Ch. 18 (Freeman, San Francisco, 1980).
9. Sarich, V. M. *Syst. Zool.* **18,** 286–295 (1969).
10. Sarich, V. M. & Cronin, J. E. in *Molecular Anthropology* (eds Goodman, M. & Tashian, R. E.) 141–170 (Plenum, New York, 1976).
11. Sarich, V. M. *Nature* **265,** 24–28 (1977).
12. Sarich, V. M. *Nature* **245,** 218–220 (1973).
13. Prager, E. M., Wilson, A. C., Lowenstein, J. M. & Sarich, V. M. *Science* **209,** 287–289 (1980).
14. Coppens, Y., Maglio, V. J., Madden, C. T. & Beden, M. in *Evolution of African Mammals* (eds Maglio, V. J. & Cooke, H. B. S.) 336–367 (Harvard University Press, 1978).
15. Bensley, B. A. *Trans. Linn. Soc. London* (*Zool.*) **9,** 83–217 (1903).
16. Sinclair, W. J. *Rep. Princeton Univ. exped. Patagonia* **4,** 333–460 (1906).
17. Archer, M. *J. Linn. Soc.* **59**(3), 217–322 (1976).
18. Archer, M. *Aust. J. Zool. Suppl.* Ser. **39,** 1–34 (1976).
19. Simpson, G. G. *Am. Mus. Novit.* **1118,** 1–6 (1941).
20. Marshall, L. G. *Syst. Zool.* **26,** 410–425 (1977).
21. Maxson, L. R., Sarich, V. M. & Wilson, A. C. *Nature* **255,** 397–400 (1975.)

Intense exercise, bone structure and blood calcium levels in vertebrates

John A. Ruben

Department of Zoology, Oregon State University, Corvallis,
Oregon 97331, USA

Albert F. Bennett

School of Biological Sciences, University of California, Irvine,
California 92717, USA

Sudden, large-scale fluctuations of systemic calcium concentrations are believed to be harmful to vertebrates. Consequently, it is generally accepted that vertebrates relatively closely monitor and regulate the calcium levels in blood and other tissues[1], often by endocrinological mechanisms which act over long time periods to promote calcium deposition on or dissolution from bone[2,3]. However, after intense bouts of maximal activity, virtually all vertebrates experience a significant reduction in blood *p*H as a result of concomitant increases in blood P_{CO_2} and lactate concentrations[4,5], and this might be expected to have wide-ranging physiological effects. We now present evidence that (1) blood plasma calcium concentrations rise abruptly and significantly as a result of intense muscular activity in all vertebrates investigated which have osseous skeletons, but not those with cartilaginous skeletons, and (2) the source of the excess calcium is bone.

We examined changes in plasma calcium concentrations and blood *p*H following bouts of intense activity in a variety of vertebrates of differing skeletal composition. Animals were stimulated to maximal activity until exhausted (~5 min) by chasing them by hand. Blood samples were taken from these individuals 10 min after cessation of activity. Similar determinations were also made on resting individuals. The results of these experiments (Table 1) indicate a significant rise in blood plasma calcium concentration after activity in all species with a bony skeleton ($P < 0.03$, *t*-test of each species) but no significant elevation in species with a cartilaginous skeleton (lamprey, dogfish; $P > 0.05$, *t*-test). Similar proportional increments in calcium concentration were also noted in whole blood samples taken before and after identical activity regimes in *Salmo*, *Squalus* and *Rattus*.

A separate set of experiments was conducted on the lizard *Iguana iguana* to determine the time course of activity-induced changes in blood *p*H, lactate and calcium concentrations. Results of these experiments (Fig. 1) indicate that, in *Iguana*, (1) the rise in plasma calcium concentration after initiation of

0028–0836/81/220411—03$01.00

activity is rapid and detectable within the first 5 min of sustainable activity, (2) the time course of elevations in plasma calcium concentration parallel temporal changes in blood *p*H and lactate concentration after initiation of activity, and (3) elevation of plasma calcium concentration as a result of intense activity may persist for 30 min or longer.

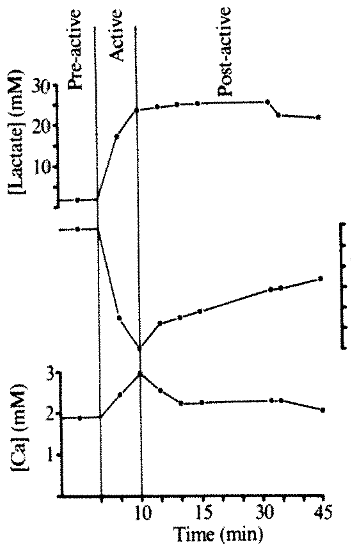


Fig. 1 Calcium and lactate concentrations and *p*H in samples of whole blood taken before, during and after 10 min of treadmill walking at 0.5 km h^{-1} in a 0.6-kg *Iguana iguana*. Three adult iguanas were exercised on a motor-driven treadmill at 35 °C. The experimental apparatus and protocol have been described elsewhere[11]. Blood was sampled before, during and after activity via an in-dwelling T-shaped cannula in the femoral artery. Anaerobically collected blood samples were analysed for *p*H as described in Table 1 legend. Additional whole blood samples were precipitated in 2× volume 0.6 M perchloric acid and analysed for [Ca] as described in Table 1 legend and [lactate] with a Boehringer–Mannheim enzymatic analysis kit[12]. Iguanas were allowed to rest for 30–60 min before pre-active blood sampling and were then exercised for 10 min at 0.5–1.0 km h^{-1}; the lizards walked steadily during the first 5 min and became fatigued and required prodding during the next 5 min. They were then allowed to recover from activity for 30–60 min. All animals experienced a rise in whole blood calcium, ranging from 0.45 to 1.15 mM, during the first 5 min of activity. Blood calcium levels peaked at 10 min activity, increasing an average of 53% above pre-active values and remained elevated for at least 30 min after exertion. Blood *p*H declined and lactate concentrations increased precipitously during activity.

What is the source of this additional plasma calcium? Major reservoirs of calcium are located in bone, scales in fish, skeletal muscle and other large organs, such as the liver. To determine the role of scales in the rise of plasma calcium in osseous piscine species, plasma calcium levels before and after activity were measured in 'scaleless' trout. These animals had scales (except those associated with the lateral line system) removed by mechanical abrasion, but remained healthy. Following a 48-h recovery period, scaleless individuals were subjected to the experimentation described in Table 1 legend. Plasma calcium concentration rose by 68% following intense activity in scaleless individuals (resting: 2.42±0.05 s.e. mM Ca, *p*H 7.52, $n = 5$; post-active: 3.99±0.11 s.e. mM Ca, *p*H 7.17, $n = 7$) the magnitude of hypercalcaemia being almost identical ($P > 0.5$, *t*-test) to that observed after activity in normally scaled individuals (Table 1). Thus, fish scales do not seem to be the major source of calcium involved in the phenomena described here.

To investigate calcium release by soft tissues, we measured calcium concentrations of skeletal muscle and liver tissue before and after activity in rainbow trout and rattlesnakes. The results (Table 2) indicate that these tissues exhibit either static or slightly increased calcium concentrations following intense activity, rather than decreasing as would be anticipated. Thus, neither can easily be assumed to be the source of the increased plasma calcium. The lack of post-active hypercalcaemia in fish with cartilaginous skeletons implies that skeletal muscle and other soft tissues in these animals do not release significant quantities of calcium even during activity-generated acidosis. It is consequently unlikely that these tissues are the source of hypercalcemia in vertebrates with osseous skeletons. Thus, bone seems to be the most likely source of this excess calcium.

The mechanism resulting in post-active hypercalcaemia might simply involve dissolution of a fraction of the crystalline calcium hydroxyapatite compartment of the osseous skeleton as a result of systemic lactic acid accumulation and *p*H depression. This hypothesis resembles that proposed for longer-term parathormone function, in which bone resorption is thought to occur, at least in part, by local secretion of lactic acid[6]. Also, the solubility of bone calcium hydroxyapatite *in vitro* has previously been shown to be sensitive to *p*H change[7].

If increased acidity of tissue fluids associated with burst activity is indeed a causal factor in bone dissolution, our data have important implications for the evolution of the vertebrate skeleton. Recent evidence[4,5] indicates that, throughout their history, vertebrates have been capable of high levels of 'burst' activity supported by anaerobic generation of lactic acid. Activity-associated dissolution of ossified skeletal tissues resul-

Table 1 Effect of intense activity on blood *p*H and plasma calcium concentration in a variety of vertebrates (mean ± s.e.)

Species	Resting			Post-active		
	*p*H	[Ca] mM	No.	*p*H	[Ca] mM	No.
Osseous skeleton						
Rainbow trout (*Salmo gairdneri*)	7.49±.04	2.31±.01	4	7.10±.03	3.97±.16	9
Longnose gar (*Lepisosteus osseus*)	7.77±.04	2.56±.07	4	7.03±.05	4.14±.05	5
White crappie (*Pomoxis annularis*)	7.93±.04	2.43±.03	4	7.42±.05	2.82±.12	6
Largemouth bass (*Micropterus salmoides*)	7.88±.04	2.49±.06	8	7.28±.05	3.01±.05	6
Rattlesnake (*Crotalus viridis*)	7.45±.02	2.58±.08	5	6.82±.07	4.08±.13	7
Laboratory rat (*Rattus rattus*)	7.50±.01	2.61±.05	4	7.33±.02	3.02±.10	6
Human (*Homo sapiens*)	7.42±.01	2.65±.02	6	7.19±.02	2.88±.04	6
Cartilaginous skeleton						
Pacific brook lamprey (*Lampetrus pacifica*)	7.91±.04	2.61±.04	6	7.23±.05	2.40±.09	8
Pacific spiny dogfish (*Squalus suckleyi*)	7.92±.05	4.31±.07	4	7.35±.02	4.54±.11	4

To determine the resting blood *p*H, 75-μl blood samples were taken anaerobically in heparinized syringes from previously undisturbed individuals of each species. Blood was collected from the ventral aorta of fish and from tail incisions in snakes and rats. Elapsed time between first handling and sample procurement was <15 s; struggling by animals was minimal during this period. The *p*H of the sample was measured immediately with a Radiometer-Copenhagen BMS 3 mark II blood microunit connected to a Radiometer-Copenhagen PHM 71 mark II acid–base analyser. The temperature of the electrode was regulated at the body temperature of the experimental subjects. Immediately after withdrawal of the previous blood sample, another sample of ~200 μl was withdrawn from each individual in a non-heparinized syringe. Plasma from these samples was then diluted at a 1:12 ratio with a calcium-suspending solvent (containing 3.60×10^{-3} M La_2O_3; 5.0×10^{-2} M HCl) and then analysed for calcium concentration on a Jarrell–Ash atomic absorption spectrophotometer equipped with a laminar flow burner. Different individuals of each species were then stimulated to maximal activity by manual prodding until exhaustion occurred (within ~5 min). After a 10-min rest period, blood samples were collected and analysed for whole blood *p*H and plasma calcium concentration as described above. Each species was maintained and exercised at an environmentally realistic temperature ($T_b = 12$ °C, *Lampetra, Salmo* and *Squalus*; $T_b = 18$ °C, *Pomoxis, Lepisosteus* and *Micropterus*; $T_b = 35$ °C, *Crotalus*). Finger prick blood samples from humans were collected before and 10 min after running exercise to complete fatigue on a staircase (~4 min). Samples were analysed as above. No significant differences in blood [Ca] were noted between individuals of different sexes in any species.

Table 2 Calcium concentration in liver and skeletal muscle before and after activity ($n = 4$ rest, 4 post-active)

Species	Liver Rest	Liver Post-active	Skeletal muscle Rest	Skeletal muscle Post-active
	Calcium concentration (mM ± s.e.)			
Rainbow trout (*Salmo gairdneri*)	0.15 ± .02	0.18 ± .03	0.31 ± .05	0.40 ± .05
Rattlesnake (*Crotalus viridis*)	0.32 ± .03	0.33 ± .05	0.37 ± .01	0.39 ± .03

Samples of liver and skeletal muscle tissue (epaxial muscle from *Salmo* and longissimus dorsi from *Crotalus*) were obtained from decapitated specimens that had been previously undisturbed or exercised to exhaustion and allowed to recover for 10 min. Samples were homogenized in a Sorvall Omnimixer, and analysed for calcium concentration by atomic absorption spectrophotometry as described in Table 1 legend.

ting in marked acute hypercalcaemia may have been a problem for vertebrates since their origin and we therefore suggest that there has been continuous selection for minimization of these effects. That selection may, at least in part, account for the evolution of several seemingly unrelated skeletal features in different groups of vertebrates. For example: (1) A primarily cartilaginous adult skeleton is acquired and maintained by several distantly related groups of vertebrates, including extant cyclostomes, elasmobranchs, chimaerans and primitive actinopterygians (Chondrostei). Although cartilage is often highly mineralized, it is virtually non-vascularized, so that the relative surface areas of cartilage exposed to post-active acidic tissue fluids must be orders of magnitude less than that for bone, which is highly vascularized. Thus, far lower rates of post-active mineral dissolution and hypercalcaemia are to be expected for cartilaginous species. (2) Acellular bone develops in extinct heterostracan ostracoderms and advanced teleost fish[8]. Acellular bone in modern teleosts is distinguished from cellular bone in other vertebrates by the absence of osteocytes and osteoclasts in the former. Consequently, like cartilage, acellular bone is considerably less vascularized than cellular bone and undoubtedly has a considerably reduced surface area exposed to post-active acidic tissue fluids. (3) An epithelia-like 'barrier' of osteocytes effectively isolates the bone-tissue fluid compartment from the tissue fluid continuum of the rest of the body. Such compartmentalization, well known in mammals[9], might serve to isolate the hydroxyapatite of bone from sudden, activity-associated pH fluctuations experienced by the soft tissues following intense activity. Each of these various developments functions, at least in part, to isolate or 'protect' most of the calcium-laden skeletal components from body tissue fluids. They have resulted in a dramatic reduction in activity-related hypercalcaemia in animals with 'protected' skeletons compared with those with apparently little or no isolation of osseous tissues from the rest of the body (Fig. 2).

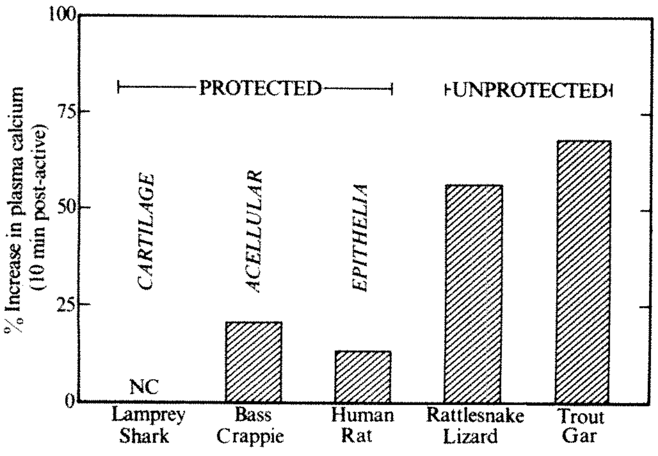


Fig. 2 Percentage increase in plasma calcium concentration after activity in vertebrates of differing skeletal type.

Thus, activity may profoundly alter blood calcium levels, even in mammals and other vertebrates with 'protected' osseous skeletons. Investigators must handle experimental subjects carefully as over-excitement may result in a rise in increments of plasma calcium concentration similar to those described here. The physiological consequences of this naturally induced, short-term hypercalcaemia have yet to be examined. However, non-activity-related acute hypercalcaemia in mammals is known to induce many deleterious physiological responses, including bradycardia, sinus arrhythmia and hypertension, disruption of normal renal function, nausea, confusion and disorientation[10]. Investigation of similar effects in vertebrates experiencing activity-related hypercalcaemia seems warranted.

This study was supported by NSF grant DEB 78-10837 (J.A.R) and grants NSF PCM 77-24208 and NIH KO4 AM00351-01 (A.F.B.). We thank T. Gleeson for assistance with blood sample analysis and T. Bradley for critical evaluation of the manuscript. We also thank N. Becker, L. Case and N. Hollingsworth for assistance and the Oregon Department of Fish and Game for providing specimens of *Salmo*, *Pomoxis* and *Micropteris*.

Received 30 July 1980; accepted 20 March 1981.

1. Phang, J. M. & Weiss, I. in *Handbook of Physiology*, Sect. 7, Vol. 7 (ed. Aurbach, G.) 157–168 (American Physiological Society, Washington DC, 1976).
2. Bentley, P. *Comparative Vertebrate Endocrinology* (Cambridge University Press, London, 1976).
3. Turner, C. & Bagnara, J. *General Endocrinology* (Saunders, Philadelphia, 1971).
4. Bennett, A. F. *A. Rev. Physiol.* **40**, 447–469 (1978).
5. Ruben, J. A. & Bennett, A. F. *Nature* **286**, 886–888 (1980).
6. Neuman, W. F., Neuman, M. W. & Meyers, C. W. *Am. J. Physiol.* **236**, 244–248 (1979).
7. McLean, F. C. & Urist, M. R. *Bone* (University of Chicago Press, 1968).
8. Hancox, N. M. *Biology of Bone* (Cambridge University Press, London, 1972).
9. Talmage, R. V. *Clin. Orthopaed.* **67**, 210–224 (1969).
10. Parfitt, A. M. & Kleerekoper, M. in *Clinical Disorders of Fluid and Electrolyte Metabolism* (eds Maxwell, M. & Kleeman, C.) 947–1151 (McGraw-Hill, San Francisco, 1980).
11. Gleeson, T. T., Mitchell, G. S. & Bennett, A. F. *Am. J. Physiol.* **239**, R174 (1980).
12. Bennett, A. F. & Licht, P. *J. comp. Physiol.* **81**, 277 (1972).

Somatic and behavioural postnatal effects of fetal injections of nerve growth factor antibodies in the rat

L. Aloe, C. Cozzari, P. Calissano & R. Levi-Montalcini

Laboratory of Cell Biology, CNR, Via Romagnosi 18/A,
00196 Rome, Italy

Previous work has shown that injections to neonatal mammals of a specific antiserum to the nerve growth factor (AS-NGF) produce irreversible destruction of sympathetic ganglia[1,2]. The animals used in these studies were mice and rats. Because this treatment does not damage other nerve cells or produce adverse effects in other tissues and organs, the immunosympathectomized young and adult rodents do not exhibit—at gross inspection—somatic or behavioural differences compared with controls except for a marked ptosis[3]. Differences become apparent, however, when the treated rodents are submitted to physiological, pharmacological and behavioural tests[4]. We report here the dramatic effects that are produced by injections of purified antibodies to NGF in rat fetuses. A preliminary account of this work has already been presented[5,6], and similar effects have been demonstrated in another laboratory.

2.5 NGF was purified by the method of Bocchini and Angeletti[7]. After two successive carboxymethyl cellulose chromatographies, NGF was run in SDS-polyacrylamide gel electrophoresis and the corresponding band cut out, eluted from the gel matrix, mixed with complete Freund's adjuvant and injected into rabbit to produce antibodies. These were purified

0028-0836/81/220413—03$01.00

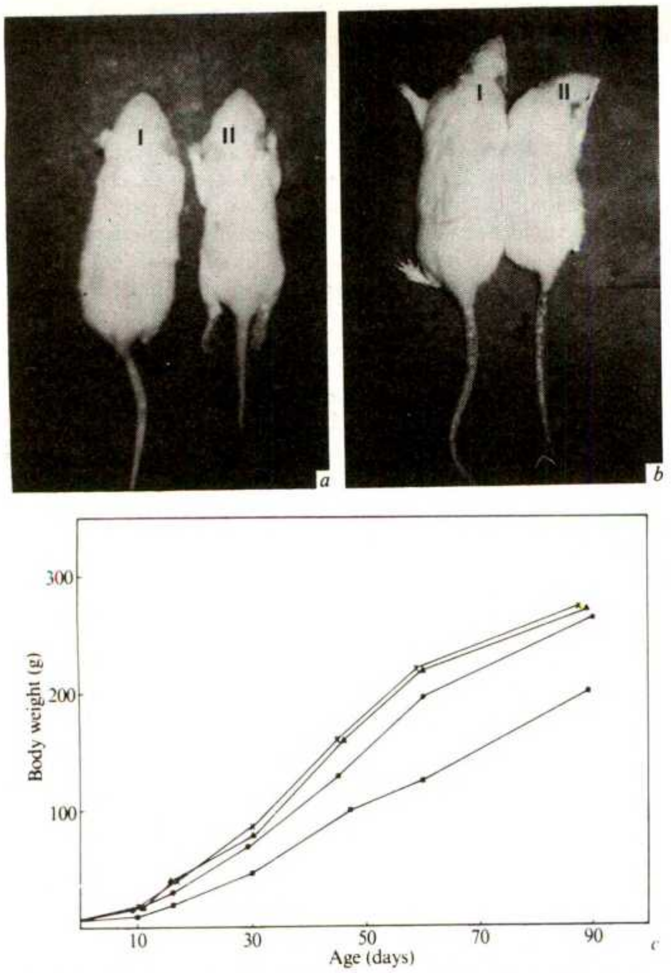


Fig. 1 Control (I) and experimental (II) rats killed respectively at 3 weeks (*a*) and 2 months (*b*) of age. Control rats were injected pre- and postnatally with 0.15 M NaCl for their first 3 weeks. Experimental rats were injected with 40 μg of purified NGF antibodies dissolved in 0.15 M NaCl. Further details are included in the text. *c*, Comparison of body weight increase in rats injected pre- and postnatally with NGF antibodies (■), 6-hydroxydopamine (●), rabbit IgG (▲) or 0.15 M NaCl (×). Each curve represents the average weight of 20 rats per group.

using the procedure of Stoeckel *et al.*[8]—affinity chromatography of the IgG fraction on Sepharose 4B with covalently linked 2.5 NGF. After absorption and washes as described, NGF antibodies were eluted with 4.5 M $MgCl_2$, dialysed extensively against twice distilled water and lyophilized. These purified preparations of NGF antibodies contained no detectable IgG fraction cross-reacting with the renin-like activity present in mouse salivary gland extract[9] and gave one single precipitin band in double-diffusion agar plates.

Female Sprague–Dawley rats were mated and at 16–17 days of gestation divided into four experimental groups. Pregnant rats were anaesthetized with ether, a middle incision made on the abdomen and the left and right uterine tubes exposed in sterile conditions. Each fetus of the first group (I) was injected subcutaneously with 40 μg of purified NGF antibodies dissolved in 20 μl of 0.15 M NaCl, using a 30-gauge needle inserted through the intact tube. The other groups served as controls and were injected with 20 μl of 0.15 M NaCl (group II), 40 μg of rabbit IgG (group III) or (group IV) with the dopamine derivative, 6-hydroxydopamine (6-OHDA), which in neonatal rodents produces massive destruction of para and pre-vertebral sympathetic ganglia[10,11]. Of the 29 pregnant rats, one of group I and one of group II died. In group I, two rats underwent abortion and a third had to have its litter delivered through caesarian intervention 2 days after the end of the gestation time. All neonatal rats of the four groups appeared vital and compared favourably at birth with untreated litters. The only noticeable difference to the untreated gestant rats was that the delivery of the litters was delayed by 24 h and in two cases by 48 h, with no apparent consequence for the mothers and the neonatal rats. The studies reported here are based on the analysis of 126 pups of the first group, 124 of the second, 48 of the third and 45 of the fourth group. In all neonatal rats the treatment was resumed immediately after birth at progressively higher doses according to the increase in body weight, as described previously[12].

Towards the end of the first postnatal week, pups injected with purified NGF antibodies (group I) were markedly smaller than those of the other three groups. 18 neonatal rats of group I died at the end of the first week and 22 before the end of the second week. In all instances death was preceded by progressive decrease in vitality and refusal to suckle even if they were taken care of by the mothers. To ensure that the weakest pups had access to nipples, the largest litters were reduced to nine pups by making available foster mothers deprived of their offspring. Six pups injected with saline (group II) and two treated with rabbit IgG (group III) died before the end of the second week. Three infants injected with 6-OHDA (group IV) died during the second postnatal week. The pups of groups II, III and IV developed normally although the latter were vital but slightly smaller than those of groups II and III (Fig. 1*c*). The 86 surviving rats treated with NGF antibodies were markedly smaller than all age-matched rats from the other three groups, as shown in Fig. 1*a–c*. Moreover, they exhibited sluggish, apathetic behaviour and moved only a few steps when exposed to strong stimuli, immediately resuming a static position, and did not show the exploratory behaviour characteristic of normal 2–4-week-old pups. On external inspection, their leg muscles appeared severely reduced in size. The behaviour of group I pups during resting of motility phases was strikingly similar to that of hypothyroid rats. Furthermore, they differed from those of other groups in their very weak reaction to pressure, prickling, burning or other noxious stimuli. Their body temperature to the hand was markedly lower than that of controls.

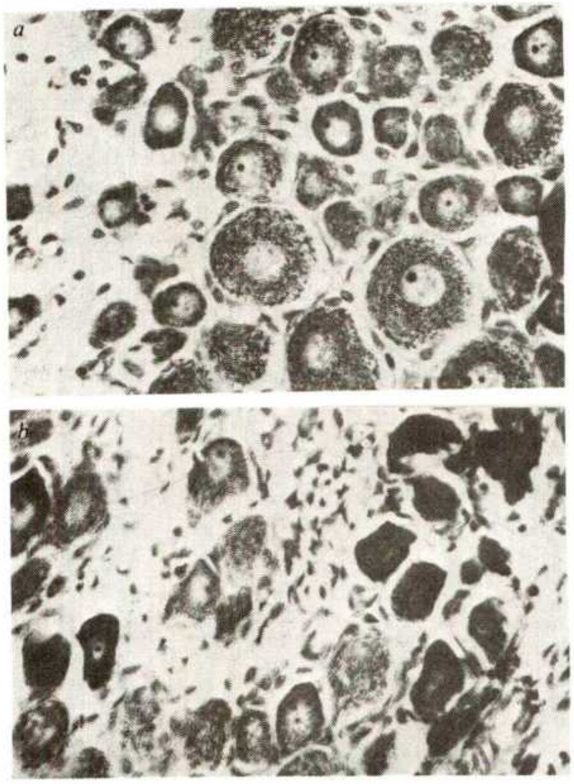


Fig. 2 Photomicrographs of sections through spinal sensory ganglia at the lumbar level of 9-day-old rats pre- and postnatally treated with 0.15 M NaCl (*a*) or NGF antibodies (*b*). Notice hypotrophic sensory nerve cells in the NGF antibody-treated rats. ×350.

Histological studies of serially sectioned sympathetic, sensory ganglia, adrenal and extra-adrenal chromaffin cells gave the following results. At birth, sympathetic ganglia were already reduced to sclerotic nodules consisting of glial cells, fibroblasts and few small atrophic sympathetic neurones. The effect was similar but more pronounced and much more precocious than that elicited by AS-NGF in neonatal rats and mice[4]. The massive degeneration of chromaffin adrenal and extra-adrenal cells has been discussed previously[13]. Sensory ganglia, which do not exhibit morphological or functional alterations in postnatal AS-NGF-injected rodents[14], appeared markedly reduced in size in 2–3-week-old group I pups. Figure 2 compares histological sections of sensory neurones of spinal ganglia of groups I and II. Nerve cells of the first group show various degrees of atrophy, ranging from decrease in cell size to evident signs of degeneration and piknosis. These findings provide evidence of severe damage to sensory neurones in rats treated with purified NGF antibodies during pre- and postnatal periods.

We also studied the accumulation of purified NGF antibodies in different organs and tissues. Fetuses, neonatal and 12-day-old rats were injected with unlabelled NGF antibodies and trace amounts of ^{125}I-NGF antibodies (1×10^6 c.p.m. per fetus or infant rat, specific activity 30,000 c.p.m. ng^{-1}). Fetuses and pups were killed 5 and 24 h later. Brain segments, hypophysis, kidneys, adrenals, salivary glands, liver, placenta, sensory and sympathetic ganglia were dissected out and the TCA-precipitable radioactivity counted. In fetuses the distribution of labelled antibodies was comparable in all organs tested, except for sympathetic and sensory ganglia, adrenal glands and hypophysis where the radioactivity was markedly higher. In a typical experiment, the specific radioactivity (counts per mg wet weight) 24 h after injections was 722, 222, 240 and 219 in sympathetic ganglia, adrenal glands, sensory ganglia and hypophysis, respectively, while the average value in all other organs, including brain regions, was 60 ± 20. In animals injected after birth, the overall level of ^{125}I-NGF antibody distribution did not differ significantly from one organ to another, except in all brain segments where the radioactivity level was about one order of magnitude lower than in animals injected before birth, probably as a result of the development of the brain–blood barrier.

Our findings suggest that neutralization of NGF or a NGF-like molecule by NGF antibodies during a critical prenatal period produces damages to sensory nerve cells, as well as to sympathetic or chromaffin cell precursors[13], and results in a syndrome reminiscent of human dysautonomic neuropathies. These findings support the hypothesis that this syndrome may result from a genetically produced NGF deficiency[15–20]. Decrease in body weight and size, and sluggish and apathetic behaviour of infant rats injected during fetal life with NGF antibodies could be due to a direct or indirect damage to one or more components of the neuroendocrine system. A NGF role in the central nervous system has been previously hypothesized[21–24]. Similar results were reported[25,26] while this paper was in preparation.

Received 7 November 1980; accepted 25 March 1981.

1. Levi-Montalcini, R. & Booder, B. *Proc. natn. Acad. Sci. U.S.A.* **42,** 384–391 (1960).
2. Cohen, S. *Proc. natn. Acad. Sci. U.S.A.* **46,** 302–311 (1960).
3. Levi-Montalcini, R. & Angeletti, P. U. *Pharmac. Rev.* **18,** 619–628 (1966).
4. Steiner, G. & Schönbaum, E. *Immunosympathectomy* (Elsevier, Amsterdam, 1972).
5. Levi-Montalcini, R. *Nerve Cells, Transmitters and Behaviour,* 84 (Elsevier, Amsterdam, 1980).
6. Levi-Montalcini, R., Aloe, L., Calissano, P. & Cozzari, C. *Abstr. 3rd Meet. int. Soc. dev. Neurosci.,* Strasbourg, 5–6 (1980).
7. Bocchini, V. & Angeletti, P. U. *Proc. natn. Acad. Sci. U.S.A.* **64,** 787–794 (1969).
8. Stoeckel, K., Gagnon, C., Guroff, G. & Thoenen, H. *J. Neurochem.* **26,** 1207–1221 (1976).
9. Cozzari, C., Angeletti, P. U., Lazar, J., Orth., H. & Cross, F. *Biochem. Pharmac.* **22,** 1321–1327 (1973).
10. Angeletti, P. U. & Levi-Montalcini, R. *Proc. natn. Acad. Sci. U.S.A.* **65,** 114–121 (1970).
11. Angeletti, P. U. & Levi-Montalcini, R. *Archs ital. Biol.* **108,** 213–221 (1970).
12. Levi-Montalcini, R., Aloe, L., Mugnaini, E., Oesch, F. & Thoenen, H. *Proc. natn. Acad. Sci. U.S.A.* **72,** 595–599 (1975).
13. Aloe, L. & Levi-Montalcini, R. *Proc. natn. Acad. Sci. U.S.A.* **76,** 1246–1250 (1979).
14. Levi-Montalcini, R. *Harvey Lecture* **60,** 217–259 (1966).
15. Riley, C. M. & Moore, R. H. *Pediatrics* **37,** 435–446 (1966).
16. Pearson, J., Pytel, B. A., Grover-Johnson, N., Axerod, F. & Dancis, J. *J. neurol. Sci.* **35,** 77–92 (1978).
17. Dolkart Gorin, P. & Johnson, E. M. *Proc. natn. Acad. Sci. U.S.A.* **76,** 5382–5386 (1979).
18. Schwartz, J. P. & Breakefield, X. O. *Proc. natn. Acad. Sci. U.S.A.* **77,** 1154–1158 (1980).
19. Otten, U., Goedert, H., Mayer, N. & Lembeck, F. *Nature* **287,** 158–159 (1980).
20. Crain, S. M., Peterson, E. R., Leibam, M. & Schulman, H. *Expl Neurol.* **67,** 205–214 (1980).
21. Freed, W. J. *Brain Res. Bull.* **1,** 393–412 (1976).
22. Benowitz, L. I. & Shashona, V. E. *Brain Res.* **172,** 561–565 (1979).
23. Walker, P., Weichsel, M. E., Fisher, D. A., Guo, S. M. & Fisher, D. A. *Science* **204,** 127–129 (1979).
24. Otten, U., Baumann, J. B. & Girard, J. *Nature* **282,** 413–414 (1979).
25. Gorin, P. D. & Johnson, E. M. *Proc. natn. Acad. Sci. U.S.A.* **76,** 5382–5386 (1980).
26. Johnson, E. M. *et al. Science* **210,** 916–918 (1980).

Mosaic distribution of opiate receptors, parafascicular projections and acetylcholinesterase in rat striatum

Miles Herkenham* & Candace B. Pert†‡

Laboratory of Neurophysiology* and Section on Biochemistry and Pharmacology, Biological Psychiatry Branch†, National Institute of Mental Health, Bethesda, Maryland 20205, USA

Although cell-stained sections have suggested heterogeneity in the neostriatum[1], more conclusive evidence has come from selective neurochemical staining for receptors, enzymes or neural connections. For example, opiate receptors in the rat striatum are concentrated in discrete islands and in a streak under the corpus callosum[2,3]—regions that lack any visible morphological distinction. In the monkey, cortical afferent fibres terminate in intricate patterns[4]. In the cat striatum, variations in the intensity of acetycholinesterase (AChE) staining[5] coincide with complex mosaic patterns of striatal efferent neurones, with cortical afferent terminations[6] and with enkephalin-like immunoreactivity[7]. In both cats[8] and monkeys[9], thalamic afferent fibres, from the parafascicular–centre median complex, terminate in discrete clusters. Even the intensely studied nigrostriatal dopamine pathway, previously assumed to project diffusely throughout the caudate, can be shown to be heterogeneous by pharmacological methods[10]. To extend this evidence, we decided to examine, on serial sections of the same rat striatum, four different histological markers of striatal heterogeneity. We now report that there is a precise mosaic 'fit' in striatal 'islands' of closely packed opiate receptors, vacancies in the termination of parafascicular projections and AChE-poor zones.

We have recently developed an *in vitro* autoradiographical method whereby cryostat-cut sections of unfixed rat brain can be attached securely to slides so that morphology is preserved, even after incubation in radiolabelled ligands to visualize drug and neurotransmitter receptors[11]. These sections are also suitable for visualizing intra-axonal transport of ^{3}H-labelled amino acids used in classical tract-tracing autoradiography[12] and for AChE enzyme histochemistry[13]. Accordingly a mixture of ^{3}H-proline and ^{3}H-leucine was injected through pipettes by iontophoresis into the thalamic parafascicular nucleus[12] of 13 Sprague–Dawley rats (weighing ~200 g). The animals were killed 10–12 days later, 11 by perfusion with 10% formol–saline and 2 by decapitation. Sections from 11 perfused rats were processed for autoradiographical visualization of labelled thalamostriatal projections as described elsewhere[12]. The two unperfused rat brains were removed and frozen in isopentane at −30°C before sectioning at −14 °C in the cryostat. Sections (25 μm thick) were pressed onto cold gelatin-coated slides and melted in place, dried at 0 °C and stored at −15 °C for several days. Alternate sections were incubated in a buffered solution at 4 °C containing ^{3}H-naloxone as described previously[11] or at 25 °C for 30 min with 1 nM ^{3}H-D-Ala-D-Leu-enkephalin in a solution containing 50 mM Tris (*p*H 7.4), 3 mM $MnCl_2$ and 2 μM GTP[14,15]. The labelled sections were then fixed in hot paraformaldehyde vapours[11], dipped in Kodak NTB-2 emulsion, exposed at −15 °C for 6–21 days and then developed for autoradiography[12]. The remaining sections were fixed in vapours,

---

‡ To whom correspondence should be addressed.

0028–0836/81/220415—04$01.00

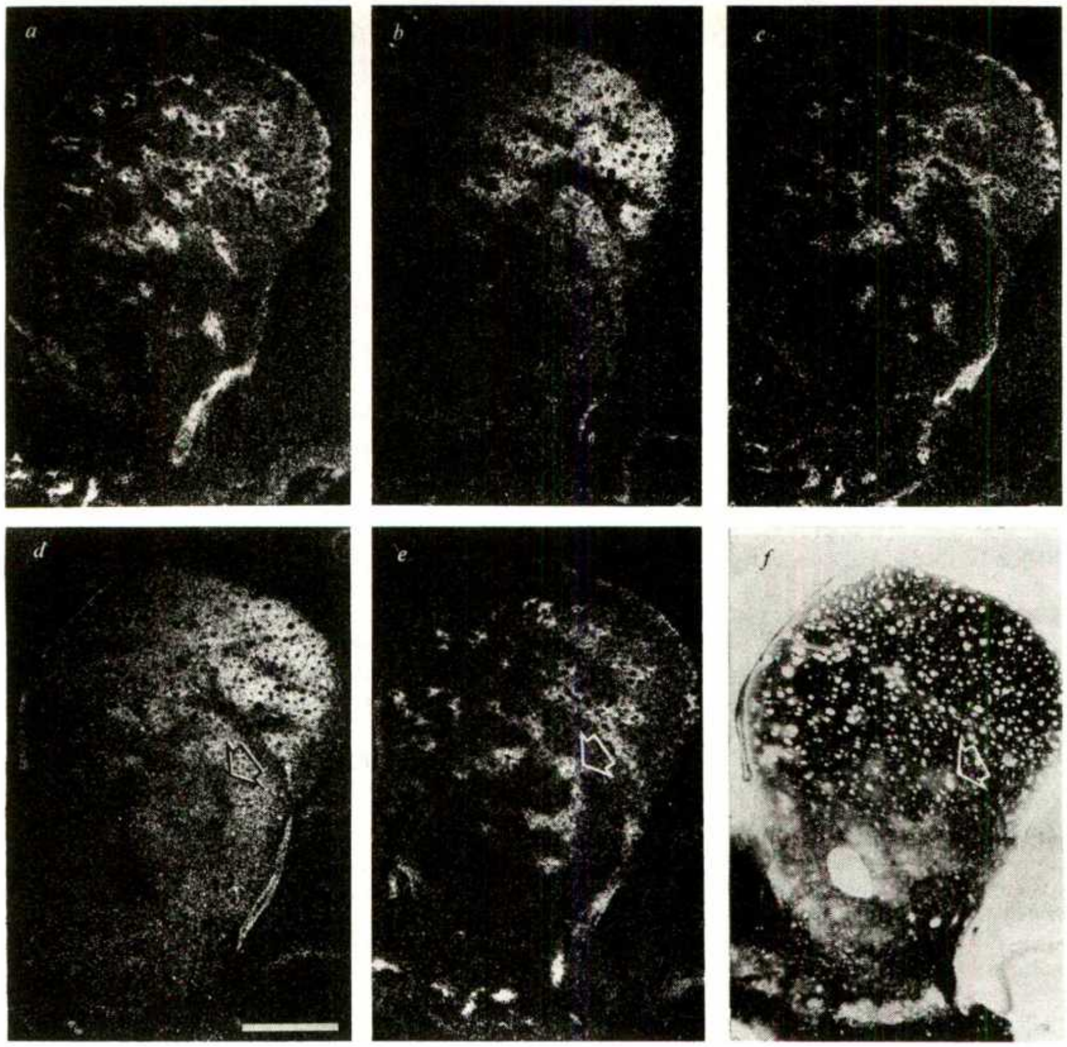


Fig. 1 Photographs of six sections taken from rostral levels of the caudate nucleus. Midline is to the left. *a, c, e* Are darkfield-illuminated autoradiograms of ^{3}H-naloxone binding visualized after the sections were exposed for 2 weeks to the emulsion. *b, d* Show the terminal distributions of thalamic parafascicular nucleus projections marked by ^{3}H-labelled amino acids in sections not treated for opiate binding but, instead, left exposed to the emulsion for 10 months. *f*, In brightfield illumination, had been stained for AChE according to a modification of the procedure described by Hardy *et al.*[13]—the sections were unfixed before AChE staining, and Iso-OMPA (K & K Laboratories, ICN) was used in place of promethazine in the incubation. Scale bar, 1 mm. Arrows point to corresponding sites exemplifying where receptor-dense zones are in register with both gaps in the thalamic termination and AChE-poor zones.

processed for autoradiography and allowed the lengthy exposure time, 6–10 months, required to visualize axonally transported label from the thalamus[12]. Occasional sections were stained for AChE[13] without prior fixation and subsequently briefly immersed in 10% formalin, stained with thionin, dehydrated and coverslipped.

Figure 1 shows that the parafascicular nucleus projections terminate throughout most of the striatum but conspicuously avoid certain irregularly shaped areas which correspond precisely to the dense opiate receptor clusters visualized in the adjacent ^{3}H-naloxone-incubated sections. As in other species[16], AChE staining in rat striatum is uneven and contains AChE-

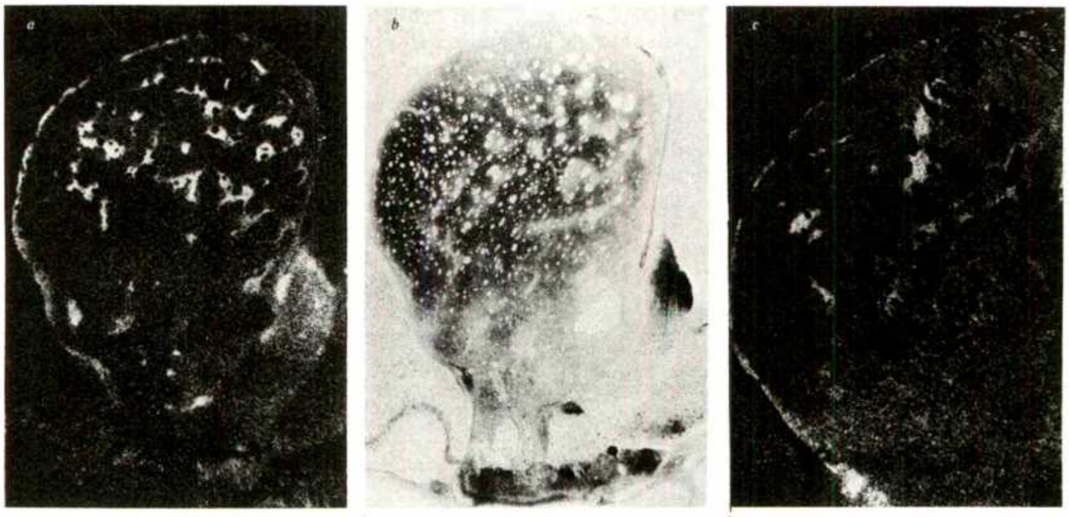


Fig. 2 *a*, Opiate receptor distribution at mid-levels of caudate. The visible streak lies immediately subjacent to the corpus callosum, seen as the dark band above. The diffuse and rather dense patch medially lies in the accumbens nucleus. A section about 200 μm removed from *a* is shown in *b*, stained for AChE. Note the similar degree of complexity and the sub-callosal streak for AChE-poor zones. In adjacent sections these would appear in register. In contrast to that pattern of correspondence, AChE-dense zones lie in the dorsomedial sector of the accumbens and in a tiny but clearly visible segment in the ventrolateral corner of the caudate nucleus—restricted zones which are also rich in opiate receptors. *c* Shows ^{3}H-naloxone binding to opiate receptors in caudal levels of striatum, where the globus pallidus and basal forebrain areas form the surrounding territories. In all three photographs medial is to the right. Magnification as in Fig. 1.

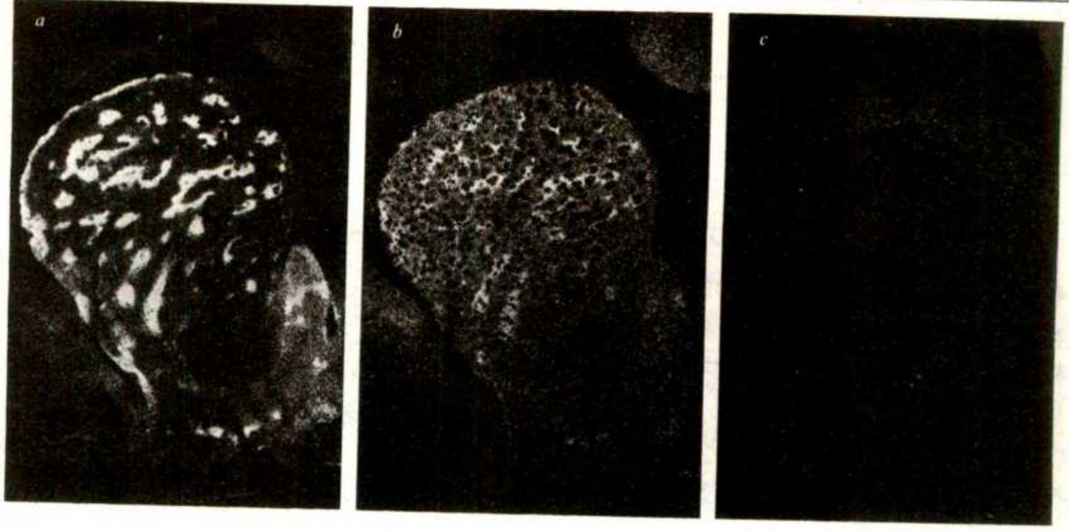


Fig. 3 *a*, *b*, Two adjacent sections incubated in ^{3}H-naloxone and ^{3}H-D-Ala-D-Leu-enkephalin, respectively. The opiate receptors labelled by the peptide are diffusely distributed throughout the striatum. When 10^{-6} M etorphine is included in either incubation medium, autoradiographical grains (white areas) are not visible. The section in *c* was incubated with 10^{-6} M etorphine and ^{3}H-D-Ala-D-Leu-enkephalin.

poor zones that coincide with striatal areas containing opiate receptor patches, although the diffuse edges and difficulty in controlling the intensity of the reaction make this concordance less dramatic. Similarly, the dense stripe of ^{3}H-naloxone-labelled opiate receptors situated just beneath the corpus callosum marks a zone that is correspondingly poor in both AChE activity and parafascicular projections. At rostral levels of striatum, ^{3}H-naloxone-labelled opiate receptors, parafascicular projections, and the AChE distribution mark corresponding zones that are rather large and diagonally orientated (Fig. 1). At mid levels of striatum, the opiate receptor distribution shows very striking borders and forms somewhat smaller patches than were visualized more rostrally (Fig. 2*a*), and the distribution of AChE is accordingly more intricate (Fig. 2*b*). In caudal striatum, the sub-callosal streak of ^{3}H-naloxone-labelled opiate receptors is still present, though narrower, and the receptor patches are smaller, often elongated and more widely separated (Fig. 2*c*). At these levels, gaps in the parafascicular projection and in AChE staining are hardly noticeable. However, wherever faint gaps can be observed they do correspond with the dense patches or stripes of receptors in adjacent sections. The receptor patches at all levels do not correlate with visible aspects of morphology, such as cell clusters[1].

In striatal sections incubated with ^{3}H-D-Ala-D-Leu-enkephalin, the labelling pattern is diffuse throughout striatal neuropil (Fig. 3) and, like ^{3}H-naloxone-labelled receptors, is absent in surrounding septum and corpus callosum. In addition, more densely labelled zones are coincident with ^{3}H-naloxone labelling observed in adjacent sections. A similar diffuse distribution of ^{125}I-D-Ala-D-Leu-enkephalin in rat striatum was recently described[17] and termed the δ receptor distribution pattern after the terminology of Lord *et al.*[18]. However, we have recently shown that ^{3}H-D-Ala-D-Leu-enkephalin can bind in the patchy as well as the diffuse pattern depending on the incubation conditions used[15], suggesting that 'μ' and 'δ' are functionall different conformations of the same, type I (ref. 14), opiate receptor.

In rat striatum, muscarinic receptor distribution[19] shows a marked heterogeneity in the neonatal animal which becomes progressively more homogeneous in the first postnatal weeks. In contrast, ^{3}H-naloxone-labelled opiate receptors are densely and diffusely distributed in the prenatal rat and form clusters in the first postnatal week, whereas ^{3}H-D-Ala-D-Leu-enkephalin-labelled receptors appear later in their homogeneous adult pattern[20]. The apparent loss of striking patterns of muscarinic receptor distribution as development proceeds may be due to the use of labelling conditions which are non-selective for muscarinic receptor subtypes[21]. With both opiate and muscarinic receptors, one receptor subtype with a clustered pattern may develop first, followed by a receptor subtype that 'fills in' to produce an apparently homogeneous adult pattern. Obviously, careful analysis of adjacent sections of developing brain for several features will be a valuable tool for analysing the sequence of development. To this end, we have found that in addition to the localization of neural connections, receptor subtypes and AChE as shown here, the peroxidase–anti-peroxidase procedure for visualizing neuropeptide-specific antibodies can be successfully performed if the sections are first briefly preincubated with 2% formalin in phosphate-buffered saline[22].

The unfixed cryostat-cut sections can also be used concurrently for visualizing catecholamine histofluorescence[23]. Several types of nigrostriatal motor behaviour in rats[24], as well as neurochemical events[25–27], have been shown to be subject to concurrent dopaminergic and 'opiatergic' modulation, even though the precise neuroanatomical substrates for these interactions remain unclear. The sub-callosal and patchy pattern of a subset of striatal dopamine terminals which develop early, fluoresce brightly and resist depletion by pharmacological manipulations[10], seems to coincide with that of μ-oid, type 1 opiate receptors. By contrast, the more homogeneous distribution of dopaminergic terminals in the adult corresponds to the more homogeneous distribution of opiate receptors visualized with ^{3}H-D-Ala-D-Leu-enkephalin (Fig.3). As brightly fluorescent dopaminergic terminals in the peripheral nervous system are found at sites where dopamine and enkephalin are co-stored and co-secreted[28], it is possible that only the subset of striatal dopaminergic terminals which seems to coincide with ^{3}H-naloxone-labelled opiate receptors contains co-stored enkephalin—one of a number of plausible notions testable by the methods and strategy described here.

We acknowledge the help of Dr R. Switzer in preparing the AChE material.

Received 17 November 1980; accepted 17 March 1981.

1. Mensah, P. L. *Brain Res.* **137,** 53–66 (1977).
2. Pert, C. B., Kuhar, M. J. & Snyder, S. H. *Life Sci.* **16,** 1849–1854 (1975).
3. Pert, C. B., Kuhar, M. J. & Snyder, S. H. *Proc. natn. Acad. Sci. U.S.A.* **73,** 3729–3733 (1976).
4. Goldman, P. S. & Nauta, W. J. H. *J. comp. Neurol.* **171,** 369–386 (1977).
5. Graybiel, A. M., Ragsdale, C. W. & Moon Edley, S. *Expl Brain Res.* **34,** 189–195 (1979).
6. Ragsdale, C. W. & Graybiel, A. M. *Soc. Neurosci. Abstr.* **5,** 78 (1979).
7. Greybiel, A. M., Ragsdale, C. W., Yoneoka, E. S. & Elde, R. P. *Soc. Neurosci. Abstr.* **6,** 342 (1980).
8. Royce, G. J. *Brain Res.* **146,** 145–150 (1978).
9. Kalil, K. *Brain Res.* **140,** 333-339 (1978).
10. Olson, L., Seiger, A. & Fuxe, K. *Brain Res.* **44,** 283–288 (1972).
11. Herkenham, M. & Pert, C. B. *Proc. natn. Acad. Sci. U.S.A.* **77,** 5532–5536 (1980).
12. Herkenham, M. *J. comp. Neurol.* **183,** 487–518 (1979).
13. Hardy, H., Heimer, L., Switzer, R. & Watkins, D. *Neurosci. Lett.* **3,** 1–5 (1976).
14. Pert, C. B., Taylor, D. P., Pert, A., Herkenham, M. A. & Kent, J. L. in *Neural Peptides and Neuronal Communication* (eds Costa, E. & Trabucchi, M.) 581–589 (Raven, New York, 1980).
15. Bowen, W. D., Gentleman, S., Herkenham, M. & Pert, C. B. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
16. Graybiel, A. M. & Ragsdale, C. W. *Proc. natn. Acad. Sci. U.S.A.* **75,** 5723–5726 (1978).
17. Goodman, R. R., Snyder, S. H., Kuhar, M. J. & Young, W. S. III *Proc. natn. Acad. Sci. U.S.A.* **77,** 6239–6234 (1980).
18. Lord, J. A. H., Waterfield, A. A., Hughes, J. & Kosterlitz, H. W. *Nature* **267,** 495–499 (1977).
19. Potter, A., Field, P. M. & Raisman, G. *Brain Res. Rev.* **1,** 185–205 (1979).
20. Kent, J. L., Pert, C. B. & Herkenham, M. *Devl Brain REF.* (in the press).
21. Birdsall, N. J. M., Hulme, E. C. & Burgen, A. *Proc. R. Soc.* A**207,** 1–12 (1980).

22. Keefer, D. A. *Cell Tissue Res.* **209,** 167-175 (1980).
23. de la Torre, J. C. & Surgeon, J. W. *Histochemistry* **49,** 81-93 (1976).
24. Pert, A. in *Characteristics and Function of Opioids Developments in Neuroscience* Vol. 4 (eds Van Ree, J. M. & Terenius, L.) 389-401 (Elsevier, Amsterdam, 1978).
25. Iwatsubo, K. & Clouet, D. H. *Biochem. Pharmac.* **24,** 1499-1503 (1975).
26. Hong, J. S., Yang, H.-Y. T., Fratta, W. & Costa, E. *J. Pharmac. exp. Ther.* **205,** 141-147 (1978).
27. Parenti, M., Gentleman, S. & Neff, N. H. *Fedn Proc.* **39,** 516 (1980).
28. Hökfelt, T. *et al.* in *Neural Peptides and Neuronal Communication* (eds Costa, E. & Trabucchi, M.) 1-23 (Raven, New York, 1980).

Cell-surface antigen distinguishes sensory and autonomic peripheral neurones from central neurones

Tom Vulliamy, Stephanie Rattray & Rhona Mirsky

MRC Neuroimmunology Project, Department of Zoology,
University College London, Gower Street, London WC1E 6BT, UK

Cell-surface markers are useful in identifying and studying different cell types in the nervous system. For example, the major cell types in rat dorsal root ganglion (DRG) cultures have been identified using a combination of tetanus toxin and antibodies to the cell-surface antigens Ran-1 and Thy-1. Neurones bind tetanus toxin and express Thy-1, Schwann cells express only Ran-1 and fibroblasts express only Thy-1 (ref. 1). A natural extension of this approach was to use antibodies to distinguish between different neuronal subpopulations on the basis of their cell-surface antigens. We therefore immunized mice with cells from rat DRG cultures, and then used the hybridoma technique of Köhler and Milstein[2] to produce monoclonal antibodies. We report here an antibody that recognizes a surface antigen which is present on all rat peripheral neurones we have studied and absent from neurones derived from the central nervous system. An antigen with reciprocal distribution, expressed only by neurones of the rat central nervous system, has been defined by Cohen and Selvendran[3].

DRG cultures from 5-day-old W/Fu or Sprague–Dawley rats were prepared as described previously[1]. Cultures were enriched for neurones by treating with 10^{-5}M cytosine arabinoside after 1 day in culture for a period of 48 h. After 2 weeks, cells were washed with Dulbecco's minimal essential medium, scraped off three Falcon flasks containing 5×10^5 neurones each, and injected intraperitoneally (i.p.) into two BALB/c mice. One week later the mice were boosted i.p. using nine flasks and a total of 4.5×10^6 neurones.

When tested by indirect immunofluorescence (IIF) 3 days after the boost, serum antibodies to neurones, Schwann cells and fibroblasts in DRG cultures could be demonstrated (for method see Fig. 1 legend). Therefore, on this day, spleen cells from one mouse were fused with a non-immunoglobulin-secreting myeloma (NS-1) using polyethylene glycol[4]. Cells were plated out into two 24-well Linbro trays at 2×10^6 cells per well in HAT (hypoxanthine–aminopterin–thymidine) medium[5]. Hybridoma clones were screened for antibody production on 5–6-day-old rat DRG cultures grown for 2–14 days on collagen-coated coverslips by IIF on living cells. Hybrids producing antibodies that bound to the surface of DRG neurones were selected and cloned twice by limiting dilution on layers of peritoneal macrophages as described elsewhere[6]. One stable cell line, 38/D7, was selected for further study. The immunoglobulin class of the antibody was shown by immunofluorescence with class-specific antibodies to be IgM.

38/D7 antibodies specifically labelled neurones and not Schwann cells or fibroblasts in DRG cultures from 5–6-day-old W/Fu, Sprague–Dawley and hooded rats. Exceptionally, in older cultures, occasional flat cells stained. When cultures were prepared from adult rats, by a method described[7] for adult mouse DRG neurones, the staining pattern was the same. Double immunofluorescence labelling showed that this antibody bound to the surface of all tetanus toxin-positive neurones in these cultures (Fig. 1*a–c*) although the intensity of staining on different cell bodies was variable. Developmentally, the antigen was first detected by double-label immunofluorescence with tetanus toxin in DRG cultures from 15-day-old rat embryos examined after 24 h *in vitro*. Cultures from 14-day-old embryos examined after 24 h *in vitro* showed no labelling.

Indirect immunofluorescence staining of primary cultures from the rat neonatal nervous system showed that peripheral neurones in the superior cervical ganglion, nodose ganglion and myenteric plexus (K. R. Jessen and R. Mirsky, unpublished observations) also bound 38/D7 antibodies. There was no staining in sciatic nerve cultures, which contain only Schwann cells and fibroblasts. Neurones in central nervous system (CNS) cultures from spinal cord (Fig. 1*d–e*), cerebellum, cerebrum and retina did not bind the antibody (Table 1). In the case of the cerebellum, cultures containing neurones were prepared from rats as old as 4 weeks. We have shown previously, using a wide variety of cell type-specific antibodies, that these cultures include Schwann cells, fibroblasts, astrocytes, oligodendroctyes, Muller cells, ependymal cells and leptomeningeal cells[1,6,8]. In addition, myoblasts and myotubes in primary muscle cultures were negative, as were melanocytes in primary cultures from the skin of newborn hooded rats. Because 38/D7 antibodies bound only to neurones of the peripheral nervous system (PNS) we conclude that the antigen is absent from all other cell types in PNS and CNS cultures, including neurones from the CNS.

Interestingly, in dissociated cultures of rat adrenal medulla (prepared as described by Unsicker *et al.*)[9] maintained in the presence of nerve growth factor (NGF), small clusters of noradrenaline-containing cells[10], with the morphology of chromaffin cells, bound 38/D7 antibodies.

In addition, the rat pheochromocytoma cell line PC12, which is derived from an adrenal medulla chromaffin cell tumour[11], bound the antibody in the absence as well as the presence of NGF in conditions where it extends neurites. Neither the rat ethylnitrosourea-induced tumour lines B35, B50, B82 and B104, which are of CNS origin[12], nor the rat–mouse glioma–neuroblastoma hybrid NG108 (ref. 13), bound antibody.

We confirmed the antigen occurrence *in situ* using 38/D7 antibodies in conjunction with peroxidase-labelled rabbit anti-mouse Ig on fixed frozen sections of DRG and cerebellum. In adult rat DRG, fixed by perfusion with 4% paraformaldehyde, there was weak labelling of the edge of neuronal cell bodies. No labelling was seen in adult cerebellar sections (Fig. 2). It is often difficult to visualize surface antigens in frozen sections[14,15]; here it was impossible to visualize 38/D7 binding in frozen sections using immunofluorescence, and difficult even with the peroxidase technique.

The relatively low levels of antibody binding seen with this monoclonal antibody in immunofluorescence and immunoperoxidase studies probably reflect low levels of antigen on the neuronal cell surface rather than a low binding constant for the

Table 1 The distribution of antigen recognized by 38/D7 antibodies in cultures of rat nervous tissue

Age of rat	Culture	Days *in vitro*	Labelled neurones
5–6 days	Dorsal root ganglion	2, 28	+
14-day embryo		1	–
15-day embryo		1	+
Adult		3	+
Newborn	Superior cervical ganglion	14	+
Newborn	Nodose ganglion	7	+
7–8 days	Myenteric plexus	8	+
5–6 days	Cerebellum	5	–
4 weeks	Cerebellum	3	–
10-day embryo	Cerebrum	3, 30	–
15-day embryo	Spinal cord	6, 16	–
Newborn	Retina	12	–

+, Antigen present; –, antigen absent.

0028-0836/81/220418—03$01.00

antibody. Concentration of the antibody supernatant 10 times or use of Ascites fluid obtained by injecting 38/D7 hybridoma cells into a BALB/c mouse did not increase the intensity of staining on neurones. Furthermore, binding did not increase when DRG cultures were fixed with 2% paraformaldehyde to stabilize the antigen–antibody complex between incubations with 38/D7 antibody and tetramethyl rhodamine-labelled goat anti-mouse Ig. The low levels of antibody binding made the use of quantitative assays and absorption studies to characterize the antigen tissue distribution difficult.

Immunofluorescence studies on PC12 cells stained in suspension showed that the binding of 38/D7 was substantially

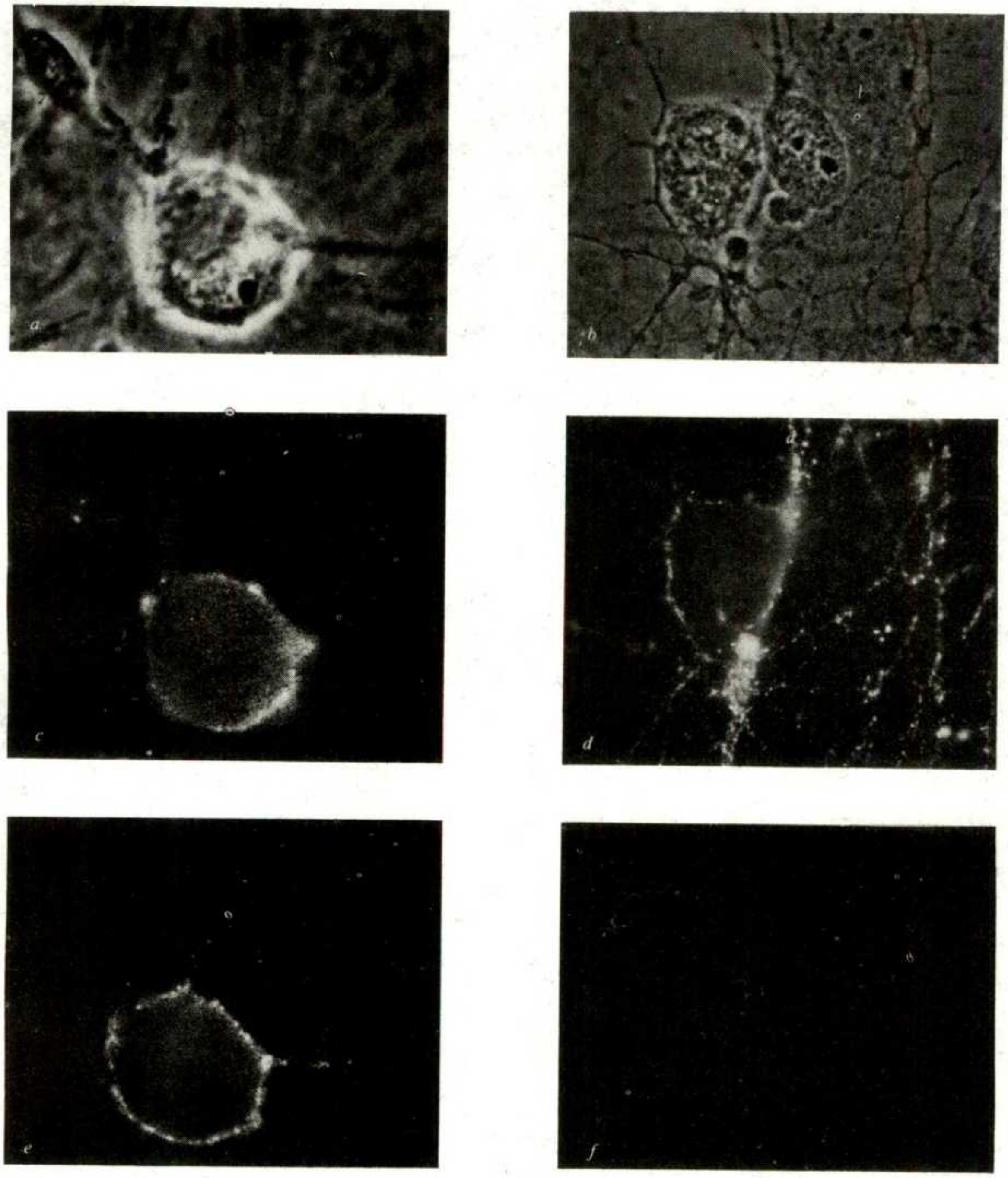


Fig. 1 38/D7 labelling of peripheral, but not central, neurones. *a*, Phase-contrast photomicrograph of a culture of rat DRG. *b*, Tetanus toxin, and *c*, 38/D7 antibodies label the neuronal cell body, as visualized by IIF; neurites are labelled, but are not in the plane of focus. Non-neuronal cells are not labelled in either case. *d*, Phase-contrast photomicrograph of a spinal cord culture. *e*, A neuronal cell body and neurites are labelled by tetanus toxin. *f*, There is no labelling by 38/D7 antibodies. Double immunofluorescent staining: cells on collagen-coated coverslips were incubated with 3–5 $\mu g\,ml^{-1}$ tetanus toxin diluted in minimal essential medium (MEM)–HEPES with 5% fetal calf serum (FCS) for 30 min at room temperature. After washing with MEM–HEPES, rabbit anti-tetanus and 38/D7 antibodies (neat hybridoma supernatant or ascites fluid (see text) diluted 1:100) were added in 40-μl volumes for 30 min at room temperature. Bound immunoglobulin was detected by further incubation for 30 min with fluorescein-conjugated goat anti-rabbit immunoglobulin (Nordic, diluted 1:50 in MEM–HEPES + 5% FCS) together with tetramethyl rhodamine-conjugated goat anti-mouse immunoglobulin (Cappell, diluted 1:100). The cells were then fixed in 5% acetic acid in ethanol for 20 min at −20 °C, and mounted in glycerol with phenylenediamine and sealed with nail varnish. The slides were examined by phase-contrast and epifluorescent illumination with either fluorescein or rhodamine filters on a Zeiss microscope. Photographs were taken with high-speed Kodak Ektachrome colour slide film. Exposure times were 30–90 s for immunofluorescence and 1/4–1/15 s for phase-contrast images. ×2,300.

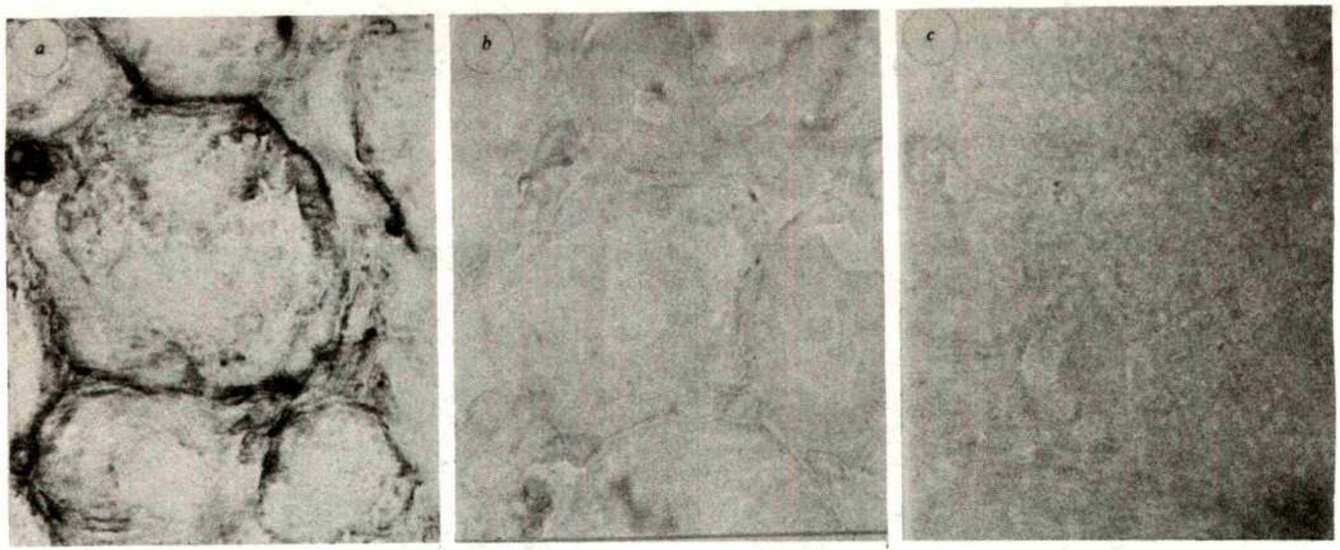


Fig. 2 38/D7 labelling of adult rat DRG neurones in section. *a*, Bright-field photomicrograph of a 10-μm adult rat DRG frozen section incubated with 38/D7 antibodies. Note precipitate visible round the edges of the neurones, faint staining within the cell. *b*, Bright-field photomicrograph of 10-μm adult rat DRG frozen section treated with A4 IgM hybridoma supernatant[3] as control. Note the general absence of staining with no distinct precipitate visible around neuronal cell bodies. *c*, Bright-field photomicrograph of a 20-μm adult rat cerebellum frozen section treated with 38/D7 antibodies. Note the general absence of staining round the Purkinje cell. Adult rat DRG were perfusion fixed with 4% paraformaldehyde, soaked overnight in 7% sucrose in phosphate-buffered saline (PBS), then frozen in liquid nitrogen. Sections (10 μm) were cut at −20 °C on a Brighton cryostat, air dried, then treated with 1% H_2O_2 in methanol for 15 min and washed in PBS. Sections were incubated with 38/D7 antibodies (neat hybridoma supernatant or ascites fluid (see text) diluted 1 : 100) or inactive hybridoma supernatant as a control for 30 min at room temperature, washed and then incubated with rabbit anti-mouse coupled to peroxidase (Nordic, diluted 1 : 100) for 30 min at room temperature. After washing in PBS and 0.05 M Tris-HCl *p*H 7.2, the sections were incubated in diaminobenzidine (0.5 mg ml^{-1}) and 0.03% H_2O_2 in 0.05 M Tris-HCl *p*H 7.2, washed, mounted in fluoromount and viewed by bright-field illumination with ×63 objective in a Zeiss Universal microscope.

reduced after preincubation of the cells at 37°C for 30 min with trypsin, suggesting that the antigen is a protein.

Finally, in a preliminary investigation of the species distribution of the antigen, DRG cultures from mouse, chicken and human showed no antibody binding, suggesting that it is rat specific.

Thus, the monoclonal antibody we have raised binds to an antigen, which is probably a protein, on the cell surface of rat peripheral neurones but does not bind to CNS neurones. The antigen is expressed by sympathetic, NGF-dependent neurones of the superior cervical ganglion as well as sensory NGF-independent neurones of the nodose ganglion and DRG. Although the origin of nodose ganglion neurones in the rat is unknown, in the chick they are of placodal origin[16], which indicates that the antigen may not be restricted to neurones derived from the neural crest. Crest-derived cells in the adrenal medulla and the pheochromocytoma PC 12 also express the antigen, whereas satellite cells and Schwann cells in various ganglia and melanocytes, which are also crest derived, do not. As the antigen is not expressed by neurones in DRG cultures until these reach an age equivalent to 16–17 days *in utero*, it can be considered a developmental antigen.

In this and the accompanying paper[3], antibodies which distinguish antigens present on the surface of neurones in the CNS and PNS are described for the first time. Tetanus toxin specifically binds to the surface of both central and peripheral neurones in culture[17], while Thy-1 is expressed on the surface of DRG neurones and on 10% or less of tetanus toxin-positive neurones in cultures of cerebellum, corpus callosum and cerebral cortex[1,8]. Nervous system-specific antigens defined by conventional antisera, which are present on the surface of neurones or neuroblastomas, have been described by several authors[18–22]. More recently, monoclonal antibodies specific to human fetal brain and various human nervous system tumours[23], to chick neurones in both the CNS and PNS[24], and to some rat retinal neurones, CNS glial cells and rat brain[25] have been described.

The two monoclonal antibodies raised in this laboratory against neurones identify different cell-surface antigens; these allow us to distinguish all CNS neurones from all PNS neurones. There may be several surface molecules that are restricted in expression either to neurones derived from the neural tube, or to neurones derived from the neural crest and placode. We find it difficult to imagine what the function of such molecules could be.

We thank the people in this neuroimmunology group, particularly Erika Abney, Jim Cohen, Jelena Gavrilovic, Mark Noble, Martin Raff and Janet Winter for help and advice, and Durward Lawson and Roy Mahoney for advice and help with photography.

Received 18 August 1980; accepted 25 March 1981.

1. Fields, K. L., Brockes, J. P., Mirsky, R. & Wendon, L. M. B. *Cell* **14,** 43–51 (1978).
2. Köhler, G. & Milstein, C. *Eur. J. Immun.* **6,** 511–519 (1976).
3. Cohen, J. & Selvendran, S. Y. *Nature* **291,** 421–423 (1981).
4. Galfre, G., Howe, S. C., Milstein, C., Butcher, G. W. & Howard, J. C. *Nature* **266,** 550–552 (1977).
5. Littlefield, J. W. *Science* **145,** 709 (1964).
6. Bartlett, P. F. *et al. Brain Res.* (in the press).
7. Scott, B. *J. Neurobiol.* **8,** 417–427 (1977).
8. Raff, M. C. *et al. Brain Res.* **174,** 283–308 (1979).
9. Unsicker, K., Krisch, B., Otten, U. & Thoenen, H. *Proc. natn. Acad. Sci. U.S.A.* **75,** 3498–3502 (1978).
10. Furness, J. B. & Costa, M. *Histochemistry* **41,** 335–352 (1975).
11. Greene, L. A. & Tischler, A. S. *Proc. natn. Acad. Sci. U.S.A.* **73,** 2424–2428 (1976).
12. Schubert, D., Heinemann, S., Carlisle, W., Tarikas, H. & Kimes, B. *Nature* **249,** 224–227 (1974).
13. Nelson, P., Clifford, C. & Nirenberg, M. *Proc. natn. Acad. Sci. U.S.A.* **73,** 123–127 (1976).
14. Moore, M. J., Dikkes, P., Reif, A. E., Romanul, F. C. A. & Sidman, R. L. *Brain Res.* **28,** 283–293 (1971).
15. Gervais, A. G. *Folio-biol., Praha* **18,** 50–52 (1972).
16. Narayanan, C. H. & Narayanan, Y. *Anat. Rec.* **196,** 71–82 (1980).
17. Mirsky, R., Wendon, L. M. B., Black, P., Stolkin, C. & Bray, C. *Brain Res.* **148,** 251–259 (1978).
18. Rutishauser, U., Thierry, J.-P., Brackenbury, R. & Edelman, G. M. *J. Cell Biol.* **79,** 371–381 (1978).
19. Rutishauser, U., Gall, W. E. & Edelman, G. M. *J. Cell Biol.* **79,** 382–393 (1978).
20. Jorgensen, O. S. & Bock, E. *J. Neurochem.* **23,** 879–880 (1974).
21. Akeson, R. & Seeger, R. C. *J. Immun.* **118,** 1995–2003 (1977).
22. Schachner, M., Ruberg, M. Z. & Carnow, T. B. *Brain Res.* B **1,** 367–377 (1976).
23. Kennett, R. H. & Gilbert, F. *Science* **203,** 1120–1121 (1979).
24. Eisenbarth, G. S., Walsh, F. S. & Nirenberg, M. *Proc. natn. Acad. Sci. U.S.A.* **76,** 4913–4917 (1979).
25. Barnstable, C. J. *Nature* **286,** 231–235 (1980).

A neuronal cell-surface antigen is found in the CNS but not in peripheral neurones

James Cohen & Siva Yogi Selvendran

MRC Neuroimmunology Project, Department of Zoology,
University College London, Gower Street, London WC1E 6BT, UK

In embryogenesis, the neural tube gives rise to the neurones of the central nervous system (CNS) whereas those in the peripheral nervous system (PNS) are generated from the neural crest and possibly also the placodes. Little is known of the part played by cell-surface molecules in this early sorting-out process and in subsequent interactions between different classes of neurone during development. Of the few known antisera that seem to define neuronal cell-surface antigens, none appears to distinguish between neuronal subpopulations[1–4]. However, the discriminating power of the immunological approach to this question has been greatly enhanced by the advent of the hybridoma technique for making monoclonal antibodies[5–8]. We have produced a hybridoma secreting an antibody that binds to the surfaces of neurones from the CNS but not from the PNS. The accompanying letter describes a monoclonal antibody with the reciprocal neuronal distribution[9].

Rat cerebellar cells prepared by trypsin digestion of tissue from 1-week-old Wistar–Furth rats as previously described[10] were grown in dissociated cell culture for 4–7 days in minimal Eagle's medium (MEM) containing 10% fetal calf serum (FCS, Gibco), 2.5% chick embryo extract (Flow) and 24 mM KCl, on poly-L-lysine (PLL, Sigma)-coated glass coverslips in 24-well trays (Linbro) or in PLL-coated plastic flasks (Falcon) and maintained in 95% air–5% CO$_2$. The cultures were enriched for neurones by growing them after 1 day in culture in the continuous presence of cytosine arabinoside (10^{-5} M, Sigma), a mitotic poison which selectively kills dividing non-neuronal cells. Previous marker studies[11] have established that cerebellar cultures grown in these conditions consist primarily of three cell types—neurones, astrocytes and fibroblasts. Cells for immunization were scraped off the culture flask in 15 mM HEPES-buffered MEM (MEMH), washed once in MEMH and 10×10^6 cells were injected intraperitoneally (without adjuvant) into female BALB/c mice. One week later the mice received a booster injection of 10×10^6 cells. When tested by indirect immunofluorescence (IIF), sera taken 24 h later were found to label all cell types in coverslip cultures of cerebellum. At the same time these were titred in an indirect radioimmunoassay (RIA) against cerebellar cells grown in the PLL-coated wells of a 96-well microtitre tray (Linbro).

Table 1 Fluorescence labelling of cells in various primary cultures with monoclonal antibody A4

Rat tissue	Neurones	Non-neuronal cells
Cerebellum	+	−
Cerebrum	+	−
Spinal cord	+	−
Hippocampus	+	−
Olfactory bulb	+	−
Retina	+	−
Dorsal root ganglion	−	−
Superior cervical ganglion	−	−
Nodose ganglion	−	−
Sciatic nerve		−
Skeletal muscle		−

Rat tissue cultures were prepared as described in the text or ref. 11 from 0–6-day-old rat tissues and grown *in vitro* for 4–6 days. Except in cases where neurones could be unambiguously identified on morphological grounds, they were recognized by double immunofluorescence assays using tetanus toxin followed by both antibody A4 and F(ab')$_2$ rabbit anti-tetanus toxin. The binding of monoclonal antibody was visualized with rhodamine-conjugated goat anti-mouse immunoglobulin and the rabbit antibody with fluoresceinated goat anti-rat immunoglobulin. Previous marker studies[11] have established that in these cultures the non-neuronal cells include astrocytes, oligodendrocytes, fibroblasts, ependymal cells, Müller cells, leptomeningeal cells, Schwann cells and microglia.

6×10^7 spleen cells from the mouse with the highest serum titre were fused 3 days after the booster immunization with 15×10^6 P3-NS1/1-Ag4-1 (NSI) cells in the presence of polyethylene glycol as described by Galfre *et al.*[12]. After fusion the cells were distributed in 7 wells of a 24-well tray (Linbro) in RPMI containing 15% FCS, sodium pyruvate (100 mM) and hypoxanthine, aminopterin and thymidine (HAT medium), fed with fresh HAT medium twice weekly for 2 weeks followed by two changes in HT medium (HAT minus aminopterin), and thereafter in RPMI/FCS.

Supernatant solutions from cultures showing vigorous hybrid cell growth (five out of seven wells) were assayed 2 weeks after the fusion by IIF and RIA against cultured cerebellar cells. Two of the five uncloned hybridoma supernatants showed relatively high binding activity (up to 10 times background) in RIAs. Only one of these, A4, will be described here.

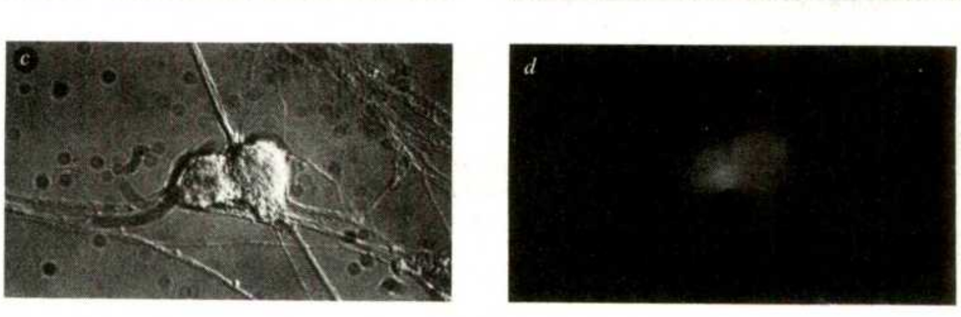


Fig. 1 A4 binding to rat CNS but not PNS neurones. Cells in cultures of 6-day-old rat cerebella (*a*, *b*) or dorsal root ganglia (*c*, *d*) were labelled after 5 days in culture with antibody A4 (1:100 dilution of immune ascites from a mouse carrying a peritoneal A4-hybridoma tumour), followed by rhodamine-conjugated goat anti-mouse immunoglobulin (G anti-MIg-Rd, Cappel; 1:100), fixed in acetic acid–ethanol (5:95) and then viewed with phase contrast (*a*) or Nomarski (*c*) and fluorescence (*b*, *d*) optics. Note that cerebellar neurones and their processes, but not the underlying fibroblastic cells, are labelled in *b* whereas the prominent sensory neurones in *d* are A4 negative. Scale bar, 10 µm.

0028–0836/81/220421—03$01.00

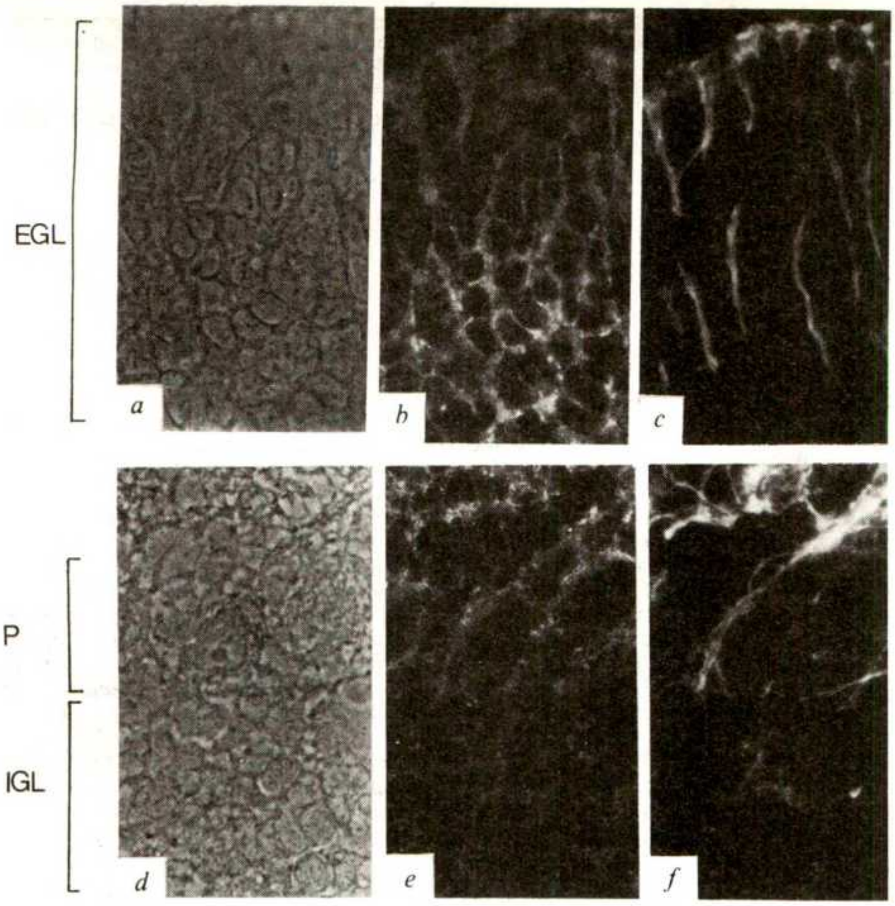


Fig. 2 Localization of A4 binding in adolescent (16-day-old) cerebellum. Small blocks of freshly dissected cerebellum were frozen in isopentane at −196 °C. Sections (6 μm) were cut on a cryostat at −15 °C and then incubated simultaneously with antibody A4 (supernatant 1:4) and rabbit anti-glial fibrillary acidic protein antiserum (RαGFAP, 1:200). After washing, binding was visualized by incubating with G anti-RIg-Rd (1:100) and sheep anti-mouse immunoglobulin fluorescein (1:100). Sections were viewed by phase (*a*, *b*) or fluorescence (*b*, *c* and *e*, *f*) optics. *b*, *e*, Antibody A4; *c*, *f*, RαGFAP; EGL, external granular layer; P, Purkinje layer; IGL, internal granular layer.

The A4 hybridoma was cloned by limiting dilution on feeder layers of mouse peritoneal cells. A4 clones continued to secrete antibody and after recloning twice, these were grown either in culture for medium collection or passaged intraperitoneally in irradiated BALB/c mice. Immune ascites from A4 tumour-bearing mice showed binding activity when diluted up to 10^7 times and was at least 100 times more potent than culture fluid. Antibody A4 was shown by Ouchterlony immunoprecipitation with class-specific reagents to be an IgM antibody.

Immunofluorescence assays of cerebellar cultures showed that antibody A4 staining was confined to the surfaces of cells with neuronal morphology (Fig. 1*a*, *b*). This neuronal staining was confirmed in double-labelling experiments with tetanus toxin, a neuronal-specific cell-surface marker[1]. A4 staining and toxin binding overlapped completely: all tetanus toxin$^+$ cells were labelled with A4 and vice versa (data not shown). The distribution of A4 binding on individual neurones was also comparable with that seen with toxin; the surfaces of cell perikarya and neurites along their entire length were stained (Fig. 1*b*).

The cell type distribution of the antigen recognized by antibody A4 was further examined by immunofluorescence labelling of cultures from a wide range of neonatal rat central and peripheral nervous tissue (Table 1). In dissociated cell cultures of spinal cord, retina, hippocampus, cerebrum and olfactory bulb, prepared as described for cerebellum[10,11], double-labelling studies showed that in all cases A4 staining and tetanus toxin binding coincided completely in cell type specificity. Thus, antibody A4 seems to recognize an antigen shared by all neurones in the CNS. By contrast, the antigen was absent from all cell types including neurones in cultures of rat dorsal root (Fig. 1*c*, *d*) sympathetic cervical and nodose ganglia. No cells stained in cultures of rat sciatic nerve, which contain Schwann cells and fibroblasts but no neurones[13] or skeletal muscle. Of the various cell lines tested, including the rat–mouse glioma–neuroblastoma hybrid NG108 (ref. 14), the ethylnitrosourea-induced rat tumour line B104 (ref. 15), and the rat phaeochromocytoma cell line PC12 (ref. 16), none bound A4.

The distribution of antigen *in vivo* was examined by IIF in 6-μm cryostat sections of immature cerebellum (postnatal day 16). Sections were double-labelled with both antibody A4 (Fig. 2*b*, *e*) and rabbit anti-glial fibrillary acidic protein (Fig. 2*c*, *f*), an intracellular glial specific marker[11], to contrast the pattern of radial glial staining with the neuronal specificity of antibody A4

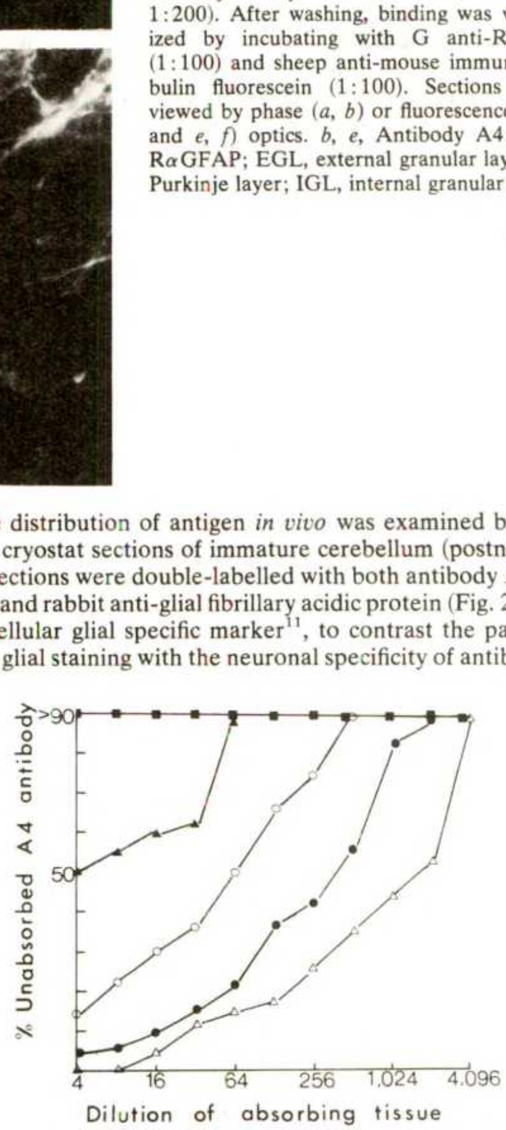


Fig. 3 Quantitative absorption analysis of A4 immune ascites with rat CNS tissue of various ages. A4 immune ascites (50 μl, 1:250) was added to 50-μl portions of doubling dilutions of 50% (v/v) tissue homogenates in MEMH and incubated for 45 min on ice. After centrifugation, the residual antibody A4 in aliquots (50 μl) of supernatant was assayed for binding to microwell cultures of 6-day-old rat cerebella (text) using ^{125}I-labelled rabbit anti-mouse μ heavy chain (50–100,000 c.p.m. per well) as second antibody. The washed microtitre plates were cut up and individual wells assayed for bound radioactivity in a γ-counter. Ordinate, residual antibody A4 expressed as a percentage of the maximal binding obtained with unabsorbed immune ascites (1:500). Each point is the mean of duplicate determinations in a single representative experiment. ■, 8-day-old whole embryo; ▲, 10-day-old embryo brain; ○, new born cerebellum; ●, 7-day-old cerebellum; △, adult cerebellum.

binding. With A4, ring fluorescence was clearly seen around most of the cell bodies in the external granular layer (Fig. 2*b*), and with lower intensity around the large Purkinje perikarya and small cells in the internal granular layer (Fig. 2*e*). No intracellular staining was seen with A4 and white matter (not shown) was only weakly positive. Control sections of rat dorsal root ganglion were unstained by this antibody (not shown, but see Fig. 2*b* in ref. 9).

The time of appearance and accretion of the antigen during rat CNS development was determined in quantitative absorption assays with homogenates of brain or cerebellum. Immune ascites fluid (1:500) was absorbed with doubling dilutions of homogenates from rats of various ages and residual antibody estimated in RIAs on cultured cerebellar cells (Fig. 3). The antigen was absent from 8-day-old fetuses but was detectable in 10-day-old embryo brain although at <1% of the adult level. By birth the concentration in brain had increased 30-fold (data not shown). Cerebellum, which develops postnatally in rodents, had 3% of adult levels at birth and 25% by day 7. The adult concentration was reached at 3 weeks (not shown).

Using the same quantitative absorption assay, several non-neural rat tissues were tested for the presence of antigen. Levels of <1% of the adult brain concentration were recorded in thymus, bone marrow, red blood cells, testes, liver and muscle (data not shown). Spleen had 3% of the adult brain level, but here binding of A4 was nonspecific. Thus, in IIF experiments with cryostat sections of adult rat spleen and brain staining with the monoclonal antibody was blocked by preincubation with normal goat serum in the case of spleen but not brain. Thus, we conclude that the antigen defined by the A4 monoclonal antibody is CNS specific.

The extent of species cross-reactivity of the antigen was examined by absorbing immune ascites with CNS tissue homogenates from chicken, frog, mouse, guinea pig, rabbit, sheep, pig, cow and human. The antigen was absent from chicken and frog and among the mammals from mouse (strains tested included BALB/c, CBA, AkR, C57 and B6) and guinea pig. However, human, rabbit, ovine, bovine and pig CNS tissue had levels of antigen comparable with those found in rat.

The exact chemical nature of the antigen remains unknown. Preliminary experiments indicate that the molecule is resistant to digestion (1 h at 37 °C) by trypsin (10 mg ml^{-1}), pronase (10 mg ml^{-1}) and neuraminidase (50 μg ml^{-1}). However, the molecule is completely destroyed by heating to 100 °C for 5 min. Deoxycholate treatment of rat brain membranes successfully solubilizes the antigen and experiments are in progress with the aim of purifying the antigen on antibody affinity columns to facilitate more detailed chemical studies.

Therefore, monoclonal antibody A4 defines a cell-surface molecule common to all CNS neurones but absent from all neurones in the PNS. Measurable amounts of the antigen are found in rat CNS as early as the 10th embryonic day and antigen rapidly accumulates in parallel with the processes of neurogenesis and neuronal maturation. In the immature cerebellum, the molecule is present on neuronal precursors including germinal cells. Thus, A4 represents a valuable new marker for studies of mammalian CNS neuronal differentiation both in culture and *in vivo*, with several advantages over tetanus toxin, the most useful cell-surface marker for neurones so far described[1]. Moreover, preliminary experiments have shown (E. Abney, B. Brown and M. Raff, unpublished observations) that antibody A4 is cytotoxic for CNS neurones in the presence of complement. This should allow exploration of the effects of selectively depleting A4$^+$ neurones on the subsequent differentiation and interactions in culture of non-neuronal cells (such as astrocytes and oligodendrocytes). The antibody should also facilitate the purification by affinity methods of A4$^+$ neurones with the aim of establishing homogeneous cultures of neurones from the CNS. The functional role of the antigen is at present obscure. An assessment of the ability of the antibody to perturb normal neuronal functions and/or differentiation, both *in vivo* and in culture, should help to elucidate this problem.

We thank Martin Raff and Stuart Bevan for helpful criticism, Erika Abney for providing rat embryo tissues, and Jack Price and Belinda Manda for technical assistance.

Note added in proof: Recent studies using cultures of optic nerve from 6-day-old rats have identified a subpopulation of A4$^+$/GFAP$^+$ process-bearing cells. Interestingly, the morphology of these cells closely resembles that of the tetanus toxin$^+$/GFAP$^+$ optic nerve astrocytes described by Raff *et al.*[11].

Received 18 August 1980; accepted 21 January 1981.

1. Mirsky, R., Wendon, L. M. B., Black, P., Stolkin, C. & Bray, D. *Brain Res.* **148,** 251–259 (1978).
2. Rutishauser, U., Thierny, J.-P., Brackenburg, R. & Edelman, G. M. *J. Cell Biol.* **79,** 371–381 (1978).
3. Rutishauser, U., Gall, W. E. & Edelman, G. M. *J. Cell Biol.* **79,** 382–393 (1978).
4. Jorgensen, O. S. & Møller, M. *Brain Res.* **194,** 419–429 (1980).
5. Köhler, G. & Milstein, C. *Nature* **256,** 495–497 (1975).
6. Kennet, R. H. & Gilbert, F. *Science* **203,** 1120–1121 (1979).
7. Eisenbarth, G. S., Walsh, F. & Nirenberg, M. *Proc. natn. Acad. Sci. U.S.A.* **76,** 4913–4917 (1979).
8. Barnstable, C. J. *Nature* **286,** 231–235 (1980).
9. Vulliamy, T., Rattray, S. & Mirsky, R. *Nature* **291,** 418–420 (1981).
10. Cohen, J., Dutton, G. R., Wilkin, G. P., Wilson, J. E. & Balazs, R. J. *Neurochemistry* **23,** 899–900 (1974).
11. Raff, M. C. *et al. Brain Res.* **174,** 283–308 (1979).
12. Galfre, G., Howe, S. C., Milstein, C., Butcher, G. W. & Howard, J. C. *Nature* **266,** 550–552 (1977).
13. Brockes, J. P., Fields, K. L. & Raff, M. C. *Nature* **266,** 364–366 (1977).
14. Nelson, P., Clifford, C. & Nirenberg, M. *Proc. natn. Acad. Sci. U.S.A.* **73,** 123–127 (1976).
15. Schubert, D. *et al. Nature* **249,** 224–227 (1974).
16. Greene, L. A. & Tischler, A. S. *Proc. natn. Acad. Sci. U.S.A.* **73,** 2424–2428 (1976).

Non-random openings and concentration-dependent lifetimes of glutamate-gated channels in muscle membrane

K. A. F. Gration, J. J. Lambert, R. Ramsey & P. N. R. Usherwood

Department of Zoology, Nottingham University, University Park,
Nottingham NG7 2RD, UK

Our preliminary results[1] with a single concentration (10^{-4} M) of L-glutamate suggested that a glutamate receptor-gated channel (glutamate channel) in locust extrajunctional muscle membrane could suddenly change its kinetic properties and switch from a short to a long lifetime mode. These results were obtained from denervated muscle[2] pretreated with the lectin concanavalin A[3,4] (Con A) to block glutamate receptor desensitization. We have since extended our patch clamp studies of locust muscle to cover a range of glutamate concentrations and to include innervated muscle and muscle which has not been treated with Con A. These new studies show that the lifetime of the glutamate channel depends on the concentration of glutamate in the patch electrode, which we explain by a multi-binding site receptor model. This model, and our finding that channel openings occur non-randomly, also accounts for the apparent transitions in channel lifetime described in our earlier publication.

Single glutamate channels associated with D-receptors were recorded from the extrajunctional membrane of innervated metathoracic extensor tibiae muscle fibres in adult locusts (*Locusta migratoria*) using the extracellular patch clamp technique[1,5–10]. Fresh pipettes were used for each concentration of glutamate (10^{-6}–10^{-2} M). Recordings from Con A-treated muscle fibres were made after pretreatment with 1–2 μM Con A for ~15 min. Experiments were performed at the muscle resting potential (−55 to −60 mV) and at ~20 °C. The open–close kinetics of single channels from muscle preparations which had not been pretreated with Con A were determined over a glutamate concentration range of 7.5×10^{-5}–7.5×10^{-4} M.

Figure 1 is a frequency histogram of the number of opening events, seen at an extrajunctional site and presumably involving a single receptor ionophore complex, in successive 5-s time intervals over a recording period of 300 s. Variation in this

0028–0836/81/220423—03$01.00

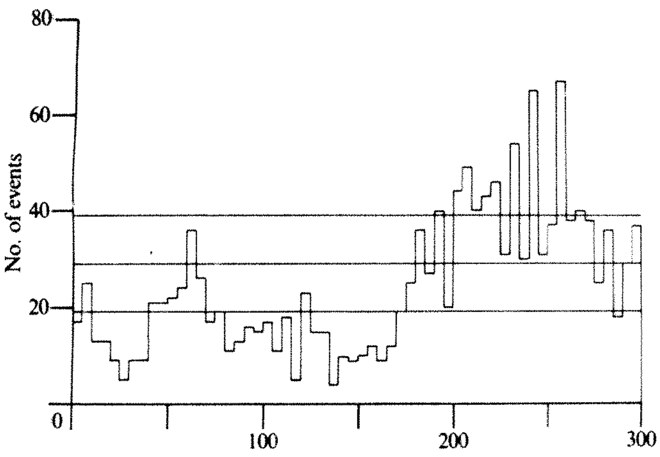


Fig. 1 Frequency histogram of the number of opening events in successive 5-ms time intervals determined from a single glutamate channel with 10^{-5} M glutamate in the patch electrode. A sample of the record from which these data were obtained is presented in Fig. 2*a*. The middle horizontal line represents the mean number of events per time bin, the upper and lower lines indicating 95% confidence limits for a Poisson distribution. Resting potential, −60 mV; temperature, ~20 °C. See text for further explanation.

frequency was a consistent feature of our recordings from Con A-treated muscles. Clearly the observed number of openings in most time bins lies outside the Poisson predicted limits. With concentrations of L-glutamate between 10^{-6} M and 10^{-4} M the intervals between openings of single glutamate channels recorded from Con A-treated preparations are usually exponentially distributed, suggesting that openings of single channels might be randomly distributed in time. However, exponential interval distributions are necessary, but not sufficient, to show that a process is Poisson and they are not very sensitive to deviations from randomness[11,12]. To determine more definitively whether the frequency of channel openings conforms to a Poisson process, the number of events occurring in equal, consecutive and non-overlapping time intervals were counted. The probability (P) of a given number of events (x) being observed in a particular interval of time (Δt) is given by:

$$P\Delta t = \frac{e^{-m} m^{x}}{x!}$$

where m = mean number of events in interval Δt. In a Poisson distribution, the mean of the distribution equals its variance (v). Mean values of the ratio v/m with, for example, 10^{-5}, 5×10^{-5} and 10^{-4} M glutamate were 2.3 ± 1.52 (s.d.), 2.38 ± 1.23 and 3.72 ± 3.6 respectively. We tested the significance of deviations of v from m for different Δts using a simple t-test:

$$t(N-1 \text{ degrees of freedom}) = \frac{(v/m)-1}{\text{s.e.}}$$

where N = number of samples and standard error (s.e.) = $(2/N-1)^{1/2}$. With concentrations of L-glutamate between 10^{-6} M and 10^{-4} M, values of v/m were all significantly ($P<0.001$) greater than the value of unity expected for a Poisson distribution. (Data from over 100 recording sites on a sample of 40 preparations studied over 15 months.) Similar results were obtained with higher concentrations of glutamate but their

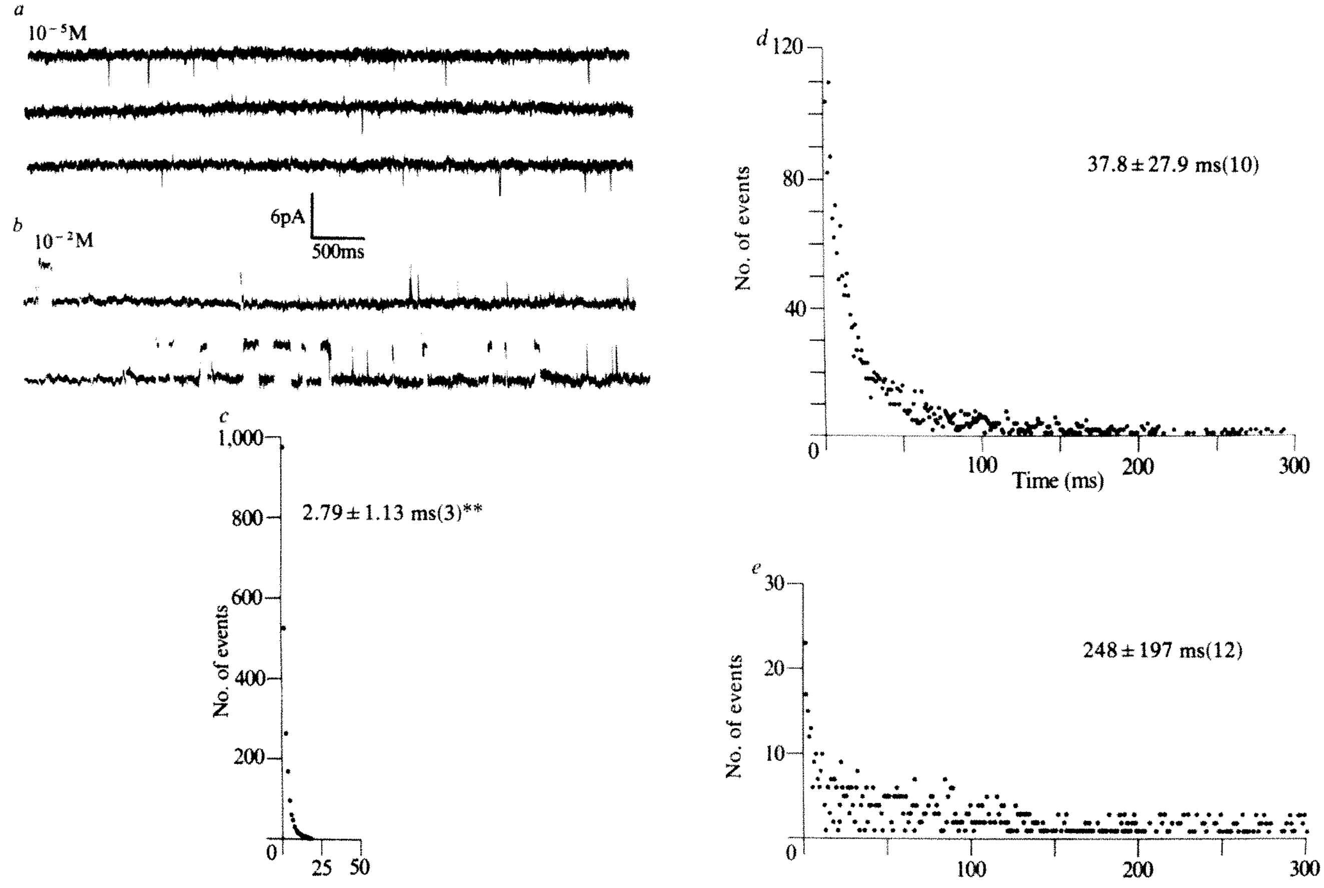


Fig. 2 Effect of glutamate concentration on channel lifetime. *a*, *b*, Single channels recorded in the presence of two different concentrations of L-glutamate, as indicated. Channel openings are downward excursions of traces for both recordings. Note channel seen with 10^{-2} M glutamate is mainly in the open state. Both records were obtained from muscle fibres which had been pretreated with Con A (10^{-6} M) to block desensitization. Resting potential, −60 mV; temperature, ~20 °C. *c*–*e*, Frequency histograms of channel lifetimes recorded in the presence of *c*, 5×10^{-5} M, *d*, 10^{-3} M, and *e*, 5×10^{-3} M glutamate. Lifetimes were measured from digitized data as the difference between the first downward deflection of the single channel current and the next upward deflection. The histograms show a progressive increase in the proportions of long open events with increasing glutamate concentration. ** Mean channel lifetime ± s.d.

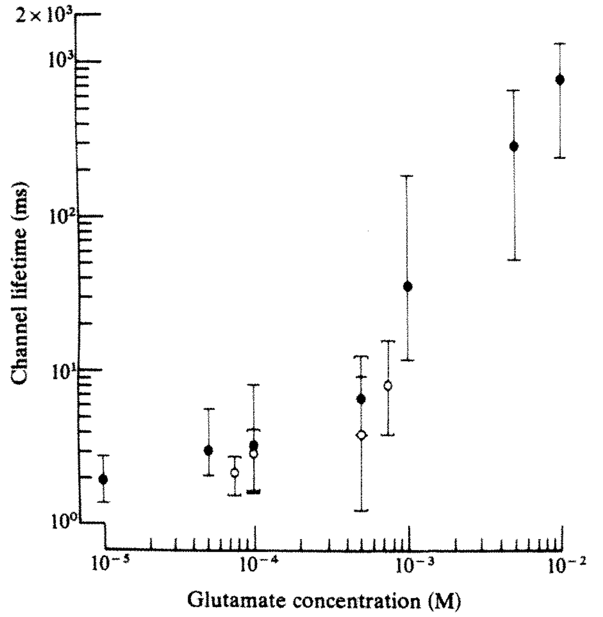


Fig. 3 Log–log plots of mean channel lifetime against glutamate concentration, obtained from Con A-pretreated (non-desensitizing) muscle preparations (●) and non-treated (desensitizing) preparations (○). Each datum point represents the average of the mean channel lifetimes of at least three recording sites (the filled circle datum point at 10^{-4} M was derived from recordings at over 50 sites). The vertical bars indicate the ranges of mean channel lifetimes for each glutamate concentration studied.

interpretation is complicated by the increased mean channel lifetime at these high concentrations. Values of v/m did not increase with increasing concentrations of glutamate between 10^{-6} M and 10^{-4} M. This suggests that non-random opening of single channels is unlikely to have resulted from unsuspected residual desensitization following treatment of the preparations with Con A, although we cannot completely exclude this possibility.

It is difficult to explain the non-random activity of single glutamate channels without invoking temporal variations either in the affinity of ligand binding site(s) or in the steps leading to ionophore gating or in the environment of the receptor–ionophore complex. Support for non-random opening of single receptor-gated channels comes from recent studies of single channel activity in planar membranes seeded with acetylcholine receptors isolated from *Torpedo* electroplax[13], where, in the presence of carbamylcholine, a slight but significant deviation from random channel opening in time was noted[14].

We have previously inferred[1,5,6] that with 10^{-4} M L-glutamate, single channel activity of 6-day denervated muscle is characterized by two or more populations of channel lifetimes, but the present studies covering a range of glutamate concentrations (10^{-6}–10^{-2} M) and using innervated rather than denervated muscles make this unlikely. At 10^{-5} M glutamate, channel lifetimes were rarely >12 ms, whereas at the other end of the concentration range studied lifetimes >1 s were observed (Fig. 2). Histograms of frequency distributions of channel lifetimes at various glutamate concentrations are illustrated in Fig. 2*c–e*, showing that the range of channel lifetimes increased with concentration of glutamate and that this was accompanied by a reduction in the proportion of channel lifetimes <10 ms. Although these trends were apparent over the entire range of concentrations studied, their impact on mean channel lifetime was not significant until the concentration exceeded ~10^{-4} M (Fig. 3). Because of the impossibility of fitting exponentials to many of the histograms we have calculated mean channel lifetimes directly from measured data. The frequency of single channel openings also increased with glutamate concentration although this parameter was obviously depressed at the higher concentrations due to the appearance of long channel lifetimes.

We have now excluded the possibility that long channel lifetimes may result from pretreatment of the muscle fibre with Con A by demonstrating their occurrence at extrajunctional sites on muscles which have not been exposed to this lectin. With glutamate concentrations ranging from 7.5×10^{-5} M to 7.5×10^{-4} M, the frequency of channel openings recorded from preparations not treated with Con A was significantly depressed compared with preparations exposed to this lectin[15] and channel openings often occurred in bursts interposed between long closed periods. Nevertheless, with the higher concentrations ($>10^{-4}$ M) of glutamate in the patch electrode this activity was characterized by long open events of the type observed at sites on Con A-treated muscles.

There remains the possibility that the long channel lifetimes result from high-frequency serial openings of normal duration separated by brief closed periods that could not be resolved due to the limited frequency response of our recording system. Our patch clamp data are usually recorded through a low-pass filter with a 3 Db cut-off frequency of 1 kHz (ref. 1) when we can reliably detect channel fluctuations of duration >300 μs at sites with good signal-to-noise ratios (~3 : 1). At the best recording sites one can resolve channel activity without introducing a low-pass filter into the recording circuit. In these conditions there is a 3-Db cut-off frequency of ~3 kHz and we can detect ~50% of channel fluctuations (openings and closings) as brief as 50 μs and 100% of fluctuations ≥100 μs duration. Our recording system was calibrated using 5-pA square pulses (50 μs – 2 ms). Even in the absence of a low-pass filter, records of channel activity obtained with high concentrations of L-glutamate in the patch electrode were still characterized by long open events.

The balance of evidence supports the contention that channel lifetime depends on the L-glutamate concentration. A receptor incorporating two or more binding sites for L-glutamate each capable of gating the receptor ionophore and with the constraint that all sites must be vacated for ionophore closure to occur would satisfactorily explain this relationship. The increased frequency of glutamate binding which would be expected to accompany an increased concentration of glutamate, would raise the probability of overlapping occupancy of the postulated binding sites on the glutamate receptor. This would result in an increase in the mean channel lifetime. The non-randomness of channel openings suggests that the affinity of the binding site might change with time. With low concentrations of agonist this non-randomness is manifested as fluctuations in the frequency of channel openings; with higher concentrations of glutamate our model would predict transient increases in the simultaneous occupancy of the binding sites and the appearance of long channel lifetimes. The model also predicts a continuous relationship between mean channel lifetime and glutamate concentration, a prediction borne out by our studies (Fig. 3).

This work was supported by SRC grants to P.N.R.U. We acknowledge the technical support of A. Tivey.

Received 15 September 1980; accepted 25 March 1981.

1. Patlak, J. B., Gration, K. A. F. & Usherwood, P. N. R. *Nature* **278**, 643–645 (1979).
2. Usherwood, P. N. R. *Nature* **223**, 411–413 (1969).
3. Mathers, D. A. & Usherwood, P. N. R. *Nature* **259**, 409–411 (1976).
4. Mathers, D. A. & Usherwood, P. N. R. *Comp. Biochem. Physiol.* **59C**, 151–155 (1978).
5. Neher, E., Sakmann, B. & Steinbach, J. H. *Pflügers Arch. ges. Physiol.* **375**, 219–228 (1978).
6. Usherwood, P. N. R., Clark, R. B., Gration, K. A. F., Ozeki, M. & Patlak, J. *J. Physiol., Paris* **75**, 615–621 (1979).
7. Clark, R. B., Gration, K. A. F. & Usherwood, P. N. R. *J. Physiol., Lond.* **301**, 60P–61P (1979).
8. Gration, K. A. F. *Neurotox. 1979 Symp. Soc. Chem. Ind.*, 169–179 (1980).
9. Gration, K. A. F., Lambert, J. J. & Usherwood, P. N. R. *Proc. 14th int. Physiol. Soc. Budapest* (in the press).
10. Usherwood, P. N. R. *INSERM Symp.* **13**, 367–383 (1980).
11. Hubbard, J. I., Llinas, R. & Quastel, D. M. J. *Electrophysiological Analysis of Synaptic Transmission* (Arnold, London, 1969).
12. Usherwood, P. N. R. *J. Physiol., Lond.* **227**, 527–551 (1972).
13. Nelson, N., Anholt, R., Linstrom, J. & Montal, M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3057–3061 (1980).
14. Schindler, H. & Quast, U. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3052–3056 (1980).
15. Gration, K. A. F., Lambert, J. J. & Usherwood, P. N. R. *J. Physiol., Lond.* **310**, 49P (1980).

Sodium channels need not open before they inactivate

Richard Horn, Joseph Patlak & Charles F. Stevens

Department of Physiology, Yale University, School of Medicine, 333 Cedar Street, New Haven, Connecticut 06510, USA

Most, if not all, types of voltage-gated channels inactivate. Channels that have been driven open by a depolarization will enter a non-conducting, inactivated state if the depolarization is maintained long enough. This inactivated state is different from the resting closed state because inactivated channels cannot be forced open by depolarization. Opinion about the mechanism of inactivation has ranged from Hodgkin and Huxley's original supposition that the inactivation process proceeds independently of events leading to channel opening[1-3] to the more recent view that a channel normally must open before inactivating[4,5]. Intermediate theories suppose that activation processes—but not necessarily the actual opening of channels—are more or less tightly coupled to inactivation[6-9]. We now provide evidence, based on recordings of currents flowing through single sodium channels, that channel opening is not a prerequisite for inactivation to occur, and that inactivation seems to proceed independently of activation.

We have used the method of recording single sodium channel currents, introduced by Sigworth and Neher[10] and modified by R.H. and J.P.[11], so that transmembrane potential is accurately known. Briefly, a fire-polished glass microelectrode with a tip pore opening about 0.5 μm in diameter is pressed on to the surface of a rat myotube grown in tissue culture conditions. When good electrical contact between the muscle surface membrane and the glass electrode is achieved, negative pressure is applied to the electrode so that the muscle membrane is drawn partly into the aperture of the electrode and forms a seal to the glass. After a good seal has developed, the electrode is quickly withdrawn from the muscle surface, and the bath solution is simultaneously changed from normal Ringer's to a solution (160 mM CsF, 5 mM Cs HEPES, *p*H 7.3) appropriate for the cytoplasmic surface of the membrane. The patch of membrane in the electrode aperture adheres to the glass, so the final arrangement is a small area of membrane separating two electrolyte-filled compartments (inside and outside the electrode). Resistance between these compartments is about $10^{10}\,\Omega$. Voltage across the isolated membrane can be controlled, and current flowing can be measured with a resolution determined by the background electrical noise of about 0.3 pA r.m.s. at a 1-kHz bandwidth. Conclusions presented here are based on observations of six different membrane patches. Individual patches usually yield useful data for between 20 and 60 min and sometimes last for >90 min. The mean open time of channels, in the conditions of the experiments described here, is about 3 ms, and the shortest openings that can be detected are about 0.15 ms. We estimate that perhaps 5% of the openings are too short to be detected. Our patches generally contained three to five channels.

We have tested the predictions made by two distinct hypotheses: first, that inactivation can only occur after a channel has opened, and second, that inactivation proceeds independently of channel opening. These two hypotheses make quite different predictions about the behaviour of channels that by chance have not opened promptly after the membrane patch was depolarized. Specifically, we have estimated the probability $P(t)$ that a channel is open t seconds after the beginning of the depolarizing stimulus, and also a conditional probability $w(T, t)$ that a channel is open at time t, given that no opening was observed for an interval of T seconds after the start of a depolarization. If channels must open before inactivation can occur, the conditional probability function should look like the overall probability of opening function, except that it should be shifted in time by T seconds; that is, $w(T, t+T) = P(t)$. The argument is that inactivation should, under our first hypothesis, proceed along a time course determined by when the channel first opens, no matter how long one happens to wait for that opening.

If inactivation occurs independently of channel opening, however, the conditional probability function should rise to a value smaller than the peak of $P(t)$—some inactivation has occurred during the T seconds' wait when the channel happened not to open—and the falling phase of the conditional probability function should coincide with that of the overall opening probability function; that is, $w(T, t) = P(t)$ for time long enough after T that channels have had a chance to open. When inactivation is not coupled to opening, the amount of inactivation present after T should be the same whether or not the channel happened to open before T. Briefly, we have found that the conditional probabilities behave as if inactivation occurred independently of channel openings.

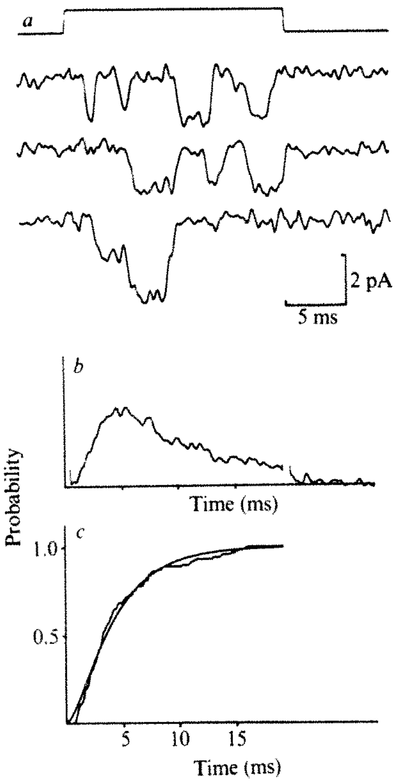


Fig. 1 Single-channel recordings in an excised patch. *a* Shows three individual records from a set of 144 depolarizations to −40 mV from a holding potential of −90 mV. The voltage pulse is depicted above the current traces. In the third trace the currents from two channels overlapped. *b* Shows an arbitrarily scaled and inverted sum of the currents from the same set of depolarizations. It estimates the probability density that a channel is open at a given time. In 27/144 traces of this series no channel openings were detected. *c* Plots the cumulative distribution of times after the voltage step to the first opening for the same set of voltage pulses. The smooth curve is a theoretical histogram for a sequential model of activation which has two closed states leading to an open state. The curve has the equation $P(t) = 1 + K_1 e^{-t/\tau_1} - K_2 e^{-t/\tau_2}$, where time constants τ_1, τ_2 = 1.26, 3.46 ms. The Ringer's solution in the pipette contained (in mM): 150 NaCl, 5 KCl, 1.5 $CaCl_2$, 1 $MgCl_2$, 5 glucose, 5 HEPES at *p*H 7.4. The cells were pretreated with 1 μg ml^{-1} α-bungarotoxin (Sigma) for 5 min before experiments to block acetylcholine receptors. Potassium channels were effectively blocked by Cs at the cytoplasmic surface. Leakage and capacity transits were subtracted from the traces by averaging the currents obtained by 25-mV hyperpolarizing voltage steps, then scaling and adding them to the currents induced by depolarization. Traces with no events were also averaged and subtracted from individual records to remove any residual capacity transients. All records were low-pass filtered at 1 kHz. Other experimental details were as described previously[11]. Temperature, 8.9 °C.

0028-0836/81/220426—02$01.00

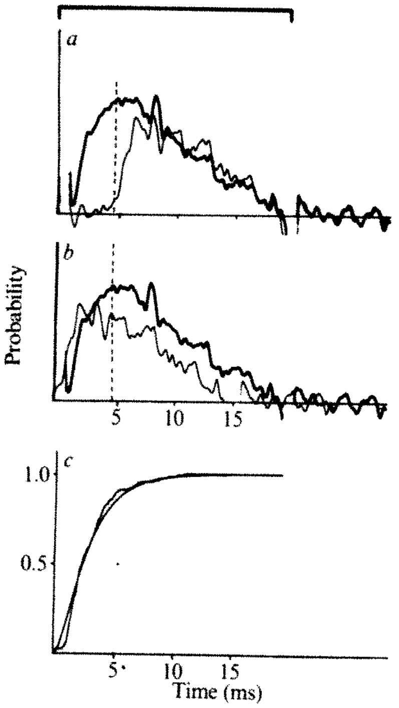


Fig. 2 Probabilities of channel opening. *a* Shows the probability (heavy line) $P(t)$ of a channel being open at a given time, as in Fig. 1*b*. Currents obtained by a 19.5-ms voltage step, indicated by the bracket at the top of the figure, to −40 mV from the holding potential of −119 mV. The records were averaged from 144 depolarizations. The patch contained three channels. The lighter line shows the average of all traces which had no openings during the first 4.55 ms, which is indicated by the dashed line. Note that the falling phase of this averaged current, which estimates the conditional probability $w(T, t)$ (see text), closely follows the falling phase of $P(t)$. *b* Replots the same data as *a*, and shows that the conditional probability, when shifted by 4.55 ms to the left, falls below the unconditional probability. *c* Plots the cumulative distribution of first latencies for this series of depolarizations. The theoretical curve has the time constants $\tau_1, \tau_2 = 0.91$, 2.18 ms; temperature 9.8 °C.

Specimen records of single depolarization epochs shown in Fig. 1*a* confirm the observations described by Sigworth and Neher[10]: the single channel currents are about 1 pA in amplitude, and exhibit the behaviour expected of a channel with an open–shut gate that behaves stochastically. When a large number of records like that shown in Fig. 1*a* are averaged, the resulting function (Fig. 1*b*) gives an estimate of the probability $P(t)$ that a channel is open at each particular time, as long as channels are independent and identical and do not change their properties over the period of observation. Simple averaging estimates this probability because all open channels contribute the same current and thus serve as counters, each of which adds a unit to the total whenever the channel is open. The distribution of times to the first opening is found as shown in Fig. 1*c*.

Conditional probabilities $w(T, t)$ were calculated, and the result, superimposed on the overall probability $P(t)$, is illustrated in Fig. 2*a* for another patch. The distribution of times until first opening for this patch is shown in Fig. 2*c*, and the conditional probabilities $w(T, t)$ were calculated by averaging all records whose first open time exceeded 4.55 ms as indicated on the histogram. After the first 4.55 ms the conditional probability rises to a peak and then joins the overall probability function $P(t)$, as predicted by the theory that supposes inactivation proceeds independently of channel opening. Clearly, in Fig. 2*a*, the amount of inactivation observed with and without the delay in openings is the same.

The first hypothesis—that inactivation proceeds only, or primarily, from the open state—predicts that $P(t)$ and conditional probability function $w(T, t)$ should have the same amplitude and general form, but that the conditional probability function should be shifted by 4.55 ms. In Fig. 2*b* the conditional probability function has been shifted to the left by 4.55 ms and superimposed on the overall opening probability function. The tail (that is, the falling phase) of the shifted conditional probability function $w(T, t+T)$ always lies below $P(t)$, contrary to what is expected if inactivation proceeds only from the open state.

The precise form of the first hypothesis (inactivation coupled to opening) we have tested is that $w(T, t+T) = P(t)$. The requirement that $w(T, t+T) = P(t)$ for values of t slightly greater than 0 tacitly assumes that the channel can exist in only one distinct closed state. The expected existence of multiple closed states is revealed, however, by changed behaviour of those channels whose opening is delayed: channels that have not opened during the first 4.55 ms have a significantly narrower distribution of times until first opening than does the entire channel population. Consequently, the conditional probability function $w(T, t)$ rises more rapidly to a peak value than does $P(t)$. This difference is not, however, large enough to produce the discrepancies in the tails of the two functions illustrated in Fig. 2*b*; nor can brief and undetected openings during the first 4.55 ms simply account for the observed departures seen in Fig. 2*b* from predictions of the first hypothesis.

Although our results are consistent with the hypothesis that inactivation proceeds independently of the activation processes, they may also be compatible with theories assuming inactivation is coupled to an activation step which precedes channel opening. If our conclusion has general applicability, however, theories which assume the inactivated state can be entered only, or predominantly, from an open state must be re-evaluated.

This work was supported by NIH grant NS 12961 to C.F.S.

Received 23 December 1980; accepted 16 March 1981.

1. Hodgkin, A. L. & Huxley, A. F. *J. Physiol., Lond.* **117**, 500–544 (1952).
2. Chandler, W. K., Hodgkin, A. L. & Meves, H. *J. Physiol., Lond.* **180**, 821–836 (1965).
3. Gillespie, J. I. & Meves, H. *J. Physiol., Lond.* **299**, 289–307 (1980).
4. Moore, L. E. & Jakobsson, E. *J. theor. Biol.* **33**, 77–89 (1971).
5. Bezanilla, F. & Armstrong, C. M. *J. gen. Physiol.* **70**, 549–566 (1977).
6. Goldman, L. & Schauf, C. L. *J. gen. Physiol.* **61**, 361–384 (1973).
7. Armstrong, C. M. & Bezanilla, F. *J. gen. Physiol.* **70**, 567–590 (1977).
8. Nonner, W. *J. Physiol., Lond.* **299**, 573–603 (1980).
9. Bean B. P. *Biophys. J.* (in the press).
10. Sigworth, F. J. & Neher, E. *Nature* **287**, 447–449 (1980).
11. Horn, R. & Patlak, J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6930–6934 (1980).

K^+ channels close more slowly in the presence of external K^+ and Rb^+

R. P. Swenson Jr & C. M. Armstrong

Department of Physiology, G4, University of Pennsylvania,
Philadelphia, Pennsylvania 19104, USA

The voltage-sensitive gates regulating the permeability of ionic channels in nerve membrane were originally thought to be insensitive to the concentration[1] and species[2,3] of readily permeant cations. In contrast, blocking cations—tetraethylammonium[4], *N*-methylstrychnine[5] or pancuronium[6]—when present in the inner end of potassium or sodium channels, hinder closing of the gates by a 'foot in the door' effect. Recent work on acetylcholine channels has clearly shown that the lifetime of the conducting state is prolonged by certain species of permeant cations[7,8]. In this case, the presence of a permeant ion near the inner end of the channel is thought to prevent closing. Somewhat similar effects have recently been reported for the K^+ channels of nerve membrane[9–11]. We report here unequivocal evidence that the permeant cations K^+ and Rb^+, when present externally, slow the closing of K^+ channels in squid axon, supporting the view that gating is indeed sensitive to the presence and species of permeant cations in the channels.

Experiments were performed on internally perfused squid giant axons, using standard axial wire voltage-clamp techniques[12]. The composition of the internal solution in all

0028-0836/81/220427—03$01.00

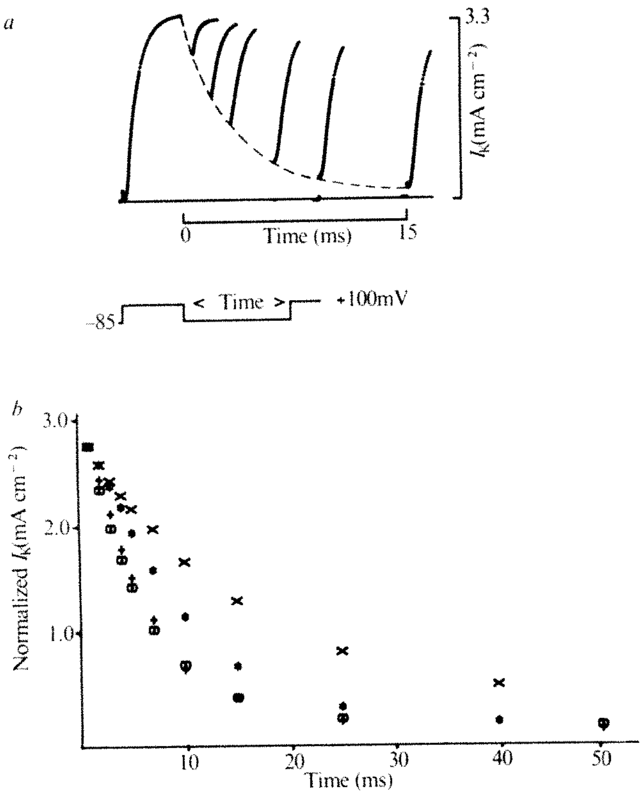


Fig. 1 Slow closing of K$^+$ channels in the presence of Rb$^+$ and K$^+$. *a*, Potassium currents recorded from a tetrodotoxin-poisoned squid axon using the double-pulse procedure described in the text. The dashed curve gives the time course of decay of the potassium conductance. *b*, The decay of K$^+$ current at −85 mV is plotted in three external solutions. The data were taken from experiments like that shown in *a*, and solutions were applied in the order listed. I_K amplitude at the end of the first pulse in the four solutions was, in order, 3.4, 3.8, 2.8 and 4.0 mA cm^{-2}. The plots have been normalized for ease of comparison. Capacity and leakage currents were subtracted electronically. All external solutions contained 2 × 10^{-7} M tetrodotoxin and their compositions were as follows: ⊡, 440 Na$^+$: 440 mM NaCl, 50 mM CaCl$_2$, 10 mM Tris-Cl (Trizma 7.0). *, 100 K$^+$, 340 Na$^+$: 100 mM KCl, 340 mM NaCl, 50 mM CaCl$_2$, 10 mM Tris-Cl. ×, 100 Rb$^+$, 340 Na$^+$: 100 mM RbCl, 340 mM NaCl, 50 mM CaCl$_2$, 10 mM Tris-Cl. +, 440 mM Na$^+$. Axon JL260A, 8 °C.

experiments was 275 mM K$^+$, 50 mM F$^+$, 225 mM glutamate, 10 mM Tris (Trizma 7.0) and 420 mM sucrose, *p*H 7.0. Compositions of the external solutions are given in Figs 1 and 2 legends. All experiments were performed at 8 °C.

The pulse protocol and associated K$^+$ currents used for measuring the rate of decay of potassium conductance (g_K) are shown in Fig. 1*a*. The K$^+$ channels were activated by a first pulse from the holding potential (usually −85 mV) to +100 mV for 5–12 ms. The potential was then restored to the holding potential for a variable interval and g_K at the end of this interval was assayed with a second pulse to +100 mV. Remaining g_K is proportional to the current at the beginning of the second step. Figure 1*a* illustrates the K$^+$ currents determined using this protocol for interpulse intervals 1–15 ms. The initial current during the second pulse becomes progressively smaller as the interval is lengthened, until after 15 ms nearly all the channels were closed. This relatively complicated protocol was used rather than simply measuring decay of the tail currents seen on return to the holding potential, because tail current kinetics can be seriously affected by depletion of ions in the confined space near the membrane[13].

Decay of the potassium conductance as determined by this protocol is plotted in Fig. 1*b*. The membrane voltage was −85 mV, and the axon was bathed successively in 440 mM Na$^+$, 100 mM K$^+$ + 340 mM Na$^+$, 100 mM Rb$^+$ + 340 Na$^+$, with a final determination in 440 mM Na$^+$ to test reversibility. The corresponding decay halftimes of the conductance were 5.2, 8.8, 14.8 and 5.6 ms. Relative to full Na$^+$ medium, equimolar substitution of 100 mM K$^+$ for Na$^+$ slows closing by a factor of 1.7, and 100 mM Rb$^+$ slows by a factor of 2.9. These effects were observed routinely, and were always reversible. This apparent stabilization of the open configuration of the channel may be visualized in terms of the foot in the door analogy, Rb$^+$ or K$^+$ playing the part of the foot. Closing of the channel is hindered or prevented by the presence of these ions at some locus in the channel.

We believe that Rb$^+$ slows closing more effectively than does K$^+$ because it has a longer dwell time in the channel, as indicated in Fig. 2, which contains plots of the instantaneous current–voltage curve for a fibre immersed in 440 mM Na$^+$, 200 mM K$^+$ + 240 mM Na$^+$, 200 mM Rb$^+$ + 240 mM Na$^+$, and again in 440 mM Na$^+$. As indicated in the inset, the K$^+$ channels were activated by a 4-ms pulse to 100 mV, and the potential was then stepped to the values given on the abscissa. Current was measured ~60 μs after the second step. With 200 mM K$^+$ the plot is nearly a straight line with a zero current intercept at +6 mV. This is somewhat more positive than the theoretical value predicted from internal and external K$^+$ concentrations because of K$^+$ ion accumulation in the Schwann cell space during the activating pulse. With 200 mM Rb$^+$, the zero current intercept is almost the same (+2 mV), but the plot has a pronounced curvature at negative potentials, showing that Rb$^+$ ions carry current inwards less efficiently than do K$^+$ ions. At large positive potentials the outward current in the presence of the two cations

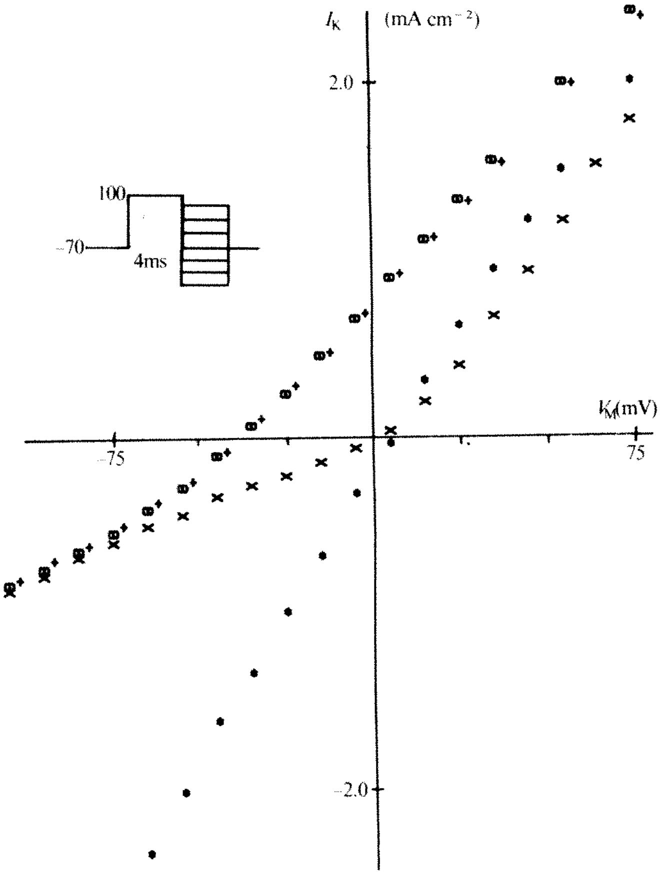


Fig. 2 Instantaneous current–voltage relations of activated K$^+$ channels in four solutions. Channels were activated by a pulse to +100 mV for 4 ms (inset), and voltage was then changed to the value given on the abscissa. Current immediately following this second voltage step was measured and is plotted as a function of voltage in the second step. Few channels can open or close in the short time before the measurement, so the curve is proportional to the *I–V* relationship of a single open K$^+$ channel. The curves were determined in three solutions in the order listed. Symbols and solutions are as described in Fig. 1 legend, except that 200 mM rather than 100 mM Rb$^+$ or K$^+$ is substituted for an equal amount of Na$^+$. Axon JL280A, 8 °C.

is nearly the same and the two curves appear to be converging. The outward current is carried by K^+ ions from the internal medium in all four plots.

The similarity of the reversal potential in both solutions indicates that Rb^+ and K^+ enter the channels equally well[14]. However, the relatively small inward current in 200 mM Rb^+ must mean that once Rb^+ is in the channel it is less mobile than K^+ and thus has a longer dwell time in the channel.

Turning again to the foot in the door, we postulate that Rb^+ slows closing more because it behaves as a stickier foot. Clearly this postulate must be tested by experiments with other permeant ions.

It is of interest to speculate where K^+ and Rb^+ are located in the channel when they exercise their gate-slowing action. There are several reasons for thinking the site is near the inner end of the channel. First, blocking cations such as tetraethylammonium are in this location when they hinder gate closing. Barium, in contrast, occupies a site near the outer end of the channel, and seems to have the opposite effect: it stabilizes the closed configuration of the channel[15]. Finally, the gate-slowing locus in acetylcholine channels is thought to be near the inner end of the channel. We hope that further examination of these questions will give a better picture of the gating machinery of K^+ channels.

Received 26 November 1980; accepted 16 March 1981.

1. Hodgkin, A. L. & Huxley, A. F. *J. Physiol., Lond.* **116,** 449–472 (1952).
2. Chandler, W. K. & Meves, H. *J. Physiol., Lond.* **180,** 788–820 (1965).
3. Hille, B. *J. gen. Physiol.* **58,** 599–619 (1971).
4. Armstrong, C. M. *J. gen. Physiol.* **58,** 413–437 (1971).
5. Cahalan, M. D. & Almers, W. *Biophys. J.* **27,** 57–74 (1979).
6. Yeh, J. Z. & Armstrong, C. M. *Nature* **273,** 387–389 (1978).
7. Ascher, P., Marty, A. & Neild, T. O. *J. Physiol., Lond.* **278,** 177–206 (1978).
8. Marchais, D. & Marty, A. *J. Physiol., Lond.* **297,** 9–45 (1979).
9. Århem, P. *Acta physiol. scand.* **108,** 7–16 (1980).
10. Dubois, J. M. & Bergman, C. *Pflugers Arch. ges Physiol.* **370,** 185–194 (1977).
11. Stühmer, W. & Conti, F. in *A. Meet. Deutsche Gesellschaft Biophysik* (eds Adam, G. & Stark, G.) (Springer, New York, 1979).
12. Bezanilla, F. & Armstrong, C. M. *J. gen. Physiol.* **70,** 549–566 (1977).
13. Frankenhauser, B. & Hodgkin, A. L. *J. Physiol., Lond.* **131,** 341–376 (1956).
14. Hille, B. in *Membranes. A Series of Advances* Vol. 3 (ed. Eisenman, G.) (Dekker, New York, 1975).
15. Armstrong, C. M. & Taylor, S. R. *Biophys. J.* **30,** 473–488 (1980).

Transcription of single-copy vitellogenin gene of *Xenopus* involves expression of middle repetitive DNA

Gerhart U. Ryffel*, Daniel B. Muellener*, Toni Wyler*, Walter Wahli† & Rudolf Weber*

* Department of Zoology, University of Bern, CH-3012 Bern, Switzerland

† Institut de Biologie Animale, Palais de Rumine, CH-1005 Lausanne, Switzerland

The genome of eukaryotes consists of single-copy DNA interspersed with middle repetitive DNA. Both types of DNA are known to be transcribed (reviewed in ref. 1). Several reports have shown that some nuclear RNA is composed of single-copy DNA transcripts covalently linked to sequences transcribed from repetitive DNA[2–4], but the significance of these nuclear transcripts is unknown; there is no convincing evidence that these nuclear transcripts containing repetitive DNA sequences are precursors to mRNAs. To elucidate this question we have analysed the structure of the two related vitellogenin genes *A1* and *A2* of *Xenopus laevis*[5,6]. The expression of these genes is initiated in liver cells by oestrogen (reviewed in refs 7–9). Preliminary experiments demonstrated that the flanking region as well as several introns of these genes contain middle repetitive DNA[9]. We have now identified and mapped six different repetitive DNA sequences within the transcribed introns of the *A1* vitellogenin gene. Further characterization of these sequences implies that they are evolutionary unstable. Analysis of nuclear RNA revealed that transcripts of the six sequences are also found in non-vitellogenic liver cells, suggesting that these repetitive DNAs are also present in transcriptional units active in the absence of oestrogen.

To identify and locate the repetitive DNA found in the *A1* vitellogenin gene we constructed subclones of the seven largest introns of the *A1* vitellogenin gene (Fig. 1). This was achieved by insertion of 400–600-base pair fragments derived from the cloned *A1* gene into the plasmid pBR322. Intron-specific clones were selected by the lack of hybridization to *A2* vitellogenin genomic clones, which are homologous in the exon sequences but differ in sequence in analogous introns[5]. To locate the origin of these intron specific subclones the labelled subclones were hybridized to blots containing genomic clones restricted with various enzymes (data not shown). In most cases this approach was sufficient to identify the intron from which the subclones were derived. The subclones were numbered according to the intron of origin and independent subclones from the same intron were distinguished by a letter. In cases where the subclones could not be mapped unambiguously we hybridized the insert of the subclone to an R-loop formed between a genomic clone and the vitellogenin mRNA. Figure 1 gives the position of 18 independent subclones derived from the 7 largest of the 33 introns. All subclones map to a unique position within *A1*, establishing that they are not repeated in other introns of the gene and they are also absent from the *A2* vitellogenin gene.

To monitor repetitive DNA sequences in the introns the subcloned DNAs were separated on an agarose gel and transferred to nitrocellulose filters. The blots were hybridized with ^{32}P-labelled total *X. laevis* DNA and the hybridized DNA was visualized by autoradiography (Fig. 2). In these hybridization conditions DNA fragments containing repetitive DNA are detected, whereas single-copy DNA does not react[9–11]. Figure 2 demonstrates that repetitive DNA sequences are present in six of the seven introns tested (see Fig. 1 for locations). As a control a slot with an *A1*-specific vitellogenin cDNA clone (pXlvc23) was included to verify that single-copy DNA did not react. The various intensities of the hybridizations suggest different repetition frequencies for the six sequences within the *Xenopus* genome. This conclusion could be confirmed by hybridizing a large excess of *Xenopus* DNA to small amounts of labelled inserts isolated from the subclones. This C_0t analysis revealed that the repetitive sequences are repeated a few hundred to a few thousand times in the genome (data not shown).

Characterization of genomic DNA has shown that in *Xenopus* as in most higher eukaryotes middle repetitive sequences are interspersed with single-copy DNA segments 1–2 kilobases long[12]. Our data demonstrate that a similar pattern of interspersion is also present within a specific gene, because the unusually large *A1* vitellogenin gene of 21 kilobases is interrupted by at least six different repetitive sequences.

To analyse the distribution of the repetitive sequences found in the *A1* vitellogenin gene throughout the *Xenopus* genome, we selected three subclones for further analysis. We isolated these cloned inserts and hybridized the labelled DNA fragments to a blot containing total *Xenopus* DNA restricted by *Eco*RI. The autoradiogram given in Fig. 3 reveals that the repetitive DNA elements tested are distributed over numerous *Eco*RI restriction fragments of different size, suggesting that the repetitive DNAs are not clustered.

Our earlier results have suggested that the *A1* and *A2* vitellogenin genes are the product of a duplication of an ancestral gene[5]. As the six repetitive sequences identified within the *A1* gene are absent in the *A2* gene, we assume that the repetitive sequences have been introduced into the *A1* gene quite recently during evolution or have diverged rapidly in the two closely related genes. To obtain further insight into the evolutionary stability of the six repetitive sequences, we hybridized a blot with all 18 intron subclones with ^{32}P-labelled total DNA from the related species *Xenopus tropicalis*. As Fig. 2 demonstrates, the repetitive sequences found in introns 13, 18 and 23 of the *A1* vitellogenin gene of *X. laevis* are not detected as repetitive DNA in *X. tropicalis*, suggesting that these sequences are evolutionary unstable. In this context it is of interest that *X. laevis* cDNA clones of the vitellogenin genes cross-react with genomic DNA of *X. tropicalis* with a specificity

0028-0836/81/220429—03$01.00

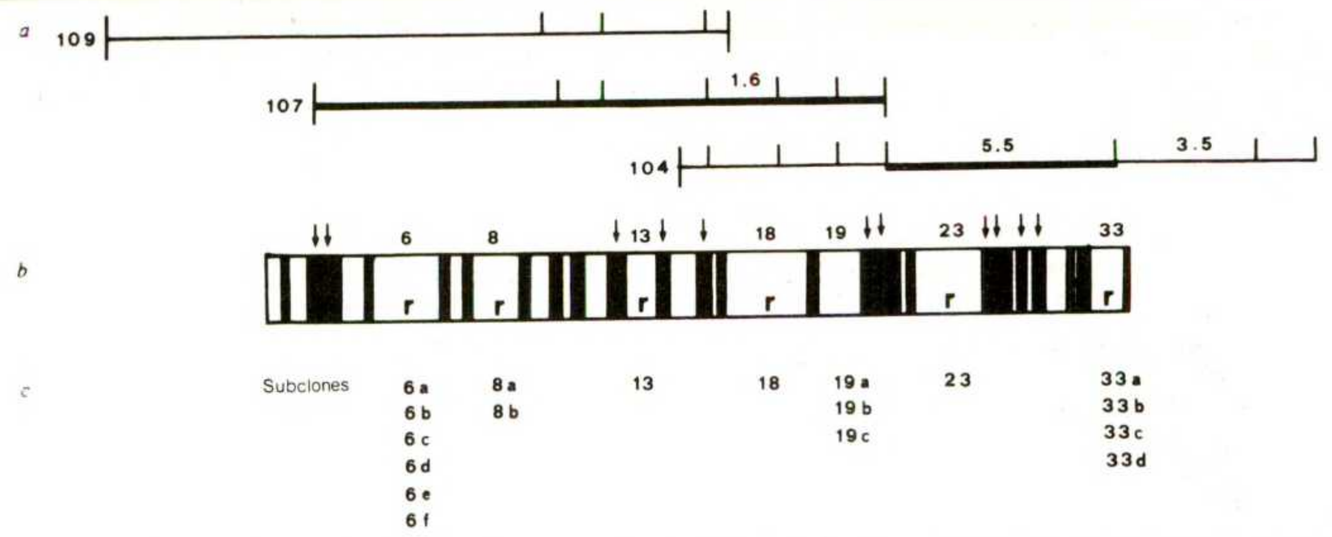


Fig. 1 Construction and location of various intron-specific subclones of the *A1* vitellogenin gene. *a*, Genomic vitellogenin gene *A1* clones with their *Eco*RI sites[5,9]. The thick line represents the restriction fragments used for subcloning and the sizes are given in kilobases. *b*, The *A1* vitellogenin gene with exons (black boxes) and introns (white boxes) is aligned to the genomic clones. The subcloned introns are numbered and small introns are indicated by arrows[5]. The letter *r* in introns denotes repetitive DNA elements identified in the present work. *c*, Location of the intron subclones obtained. The genomic clone λXlv107 and the 5.5-kilobase *Eco*RI fragment of the genomic clone λXlv104 were partially digested with *Alu*I and *Hae*III, fragments of 400–600 base pairs were isolated electrophoretically[17] and cloned into the *Pst*I site of pBR322 by G · C tailing[22]. Tetracycline-resistant, ampicillin-sensitive clones were screened by hybridization with *A1* and *A2* vitellogenin genomic clones[23]. Clones hybridizing with *A1* but not with *A2* were isolated and again tested for lack of hybridization with the *A2* genomic clone λXlv128, containing the complete *A2* gene[5]. Further details and the methods used for mapping the position of the subclones are given in the text. Subclone 18 was prepared by *Hha*I digestion of the isolated 1.6-kilobase *Eco*RI fragment of the genomic clone λXlv107 and inserting a 1-kilobase *Hha*I-*Eco*RI fragment into the *Pst*I site of pBR322 by G · C tailing. This clone also failed to react with the *A2* genomic clone λXlv 128.

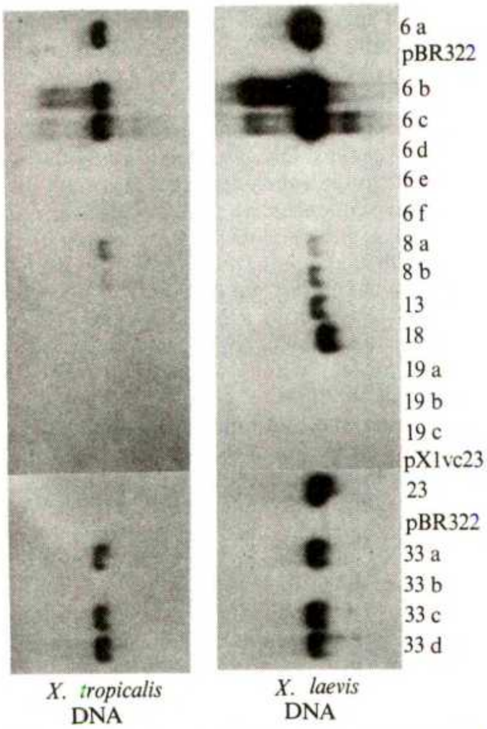


Fig. 2 Identification of repetitive sequences within the subclones. The intron subclones, the vector pBR322 and the *A1* vitellogenin cDNA clone pXlvc23[16] were digested with *Eco*RI, separated electrophoretically on a 1% agarose gel and transferred to nitrocellulose filters. The blot was hybridized[9] at 37 °C for 72 h with *X. laevis* or *X. tropicalis* DNA labelled by nick translation, washed and exposed for autoradiography as described elsewhere[9]. Electrophoresis is from right to left.

suggesting that these sequences differ by about 10% (R. Jaggi, personal communication). In complementary experiments we hybridized labelled isolated inserts of intron subclones 6, 13 and 18 to blots of *Eco*RI-digested genomic DNA. The results included in Fig. 3 confirm and extend the above conclusions. The repetitive DNA of intron 6 is present in *X. tropicalis* and is distributed over many *Eco*RI restriction fragments whereas the repetitive DNAs of the intron subclones 13 and 18 do not cross-hybridize with *X. tropicalis* DNA, suggesting that these sequences are even absent as single-copy DNA.

Analysis of nuclear RNAs containing vitellogenin mRNA sequences has shown directly that the large introns 11–33 of the vitellogenin gene are transcribed as part of a large polyadenylated precursor RNA[13]. Assuming that all the other introns are also transcribed, we conclude that the transcription of the single-copy *A1* vitellogenin gene involves the expression of a specific set of at least six repetitive DNA sequences located in introns. To answer the question of whether these repetitive sequences are also present in other transcriptional units, we analysed different nuclear RNA preparations for transcripts complementary to the 18 subclones. First, we hybridized a blot of all 18 subclones with a labelled polyadenylated nuclear RNA enriched for the vitellogenin precursor RNA. As Fig. 4 illustrates, all intron subclones hybridize with this RNA preparation, which agrees with the assumption that indeed all introns are transcribed as part of the primary transcript. If we used total polyadenylated nuclear RNA from an oestrogen-treated female producing large amounts of vitellogenin, some subcloned intron sequences could hardly be detected in these nuclear RNAs, indicating that these sequences are too dilute in this RNA preparation (Fig. 4). To elucidate whether the repetitive sequences found in the *A1* vitellogenin gene are also part of other transcriptional units, we used labelled polyadenylated nuclear RNA from the liver of an unstimulated male, known not to produce any vitellogenin mRNA[9,14,15]. Control experiments included in Fig. 4 confirm the absence of vitellogenin gene transcripts as shown by the failure of this RNA to hybridize with the cloned vitellogenin cDNA pXlvc23. On the other hand, the data given in Fig. 4 clearly establish that this RNA preparation contains sequences hybridizing to all six repetitive sequences found in the *A1* vitellogenin gene. This implies that repetitive sequences of *A1* are also found in transcriptional units active even in the absence of oestrogen.

Our analysis of the *A1* vitellogenin gene of *X. laevis* leads to the following main observations on transcribed middle repetitive DNA sequences.

As the six repetitive DNA sequences found in the *A1* gene are absent from the closely related *A2* gene, and as these two genes are apparently under coordinate control by oestrogen[16,17], one is tempted to conclude that these repetitive sequences have no function for gene expression.

The repetitive sequences of the *A1* gene have certain characteristics of mobile genetic elements. Assuming that genes *A1* and *A2* have arisen by gene duplication, the dramatic change in sequence and length of analogous introns[5] suggests that these sequences may have evolved to a large degree by insertions and deletions. Apparently, the inserted or deleted DNA sequences are frequently repetitive DNA. This holds also for the *A1* vitellogenin gene variant which contains a repetitive sequence in the intron 11 (ref. 9).

Comparison of the repetitive DNAs of the *A1* vitellogenin gene with those of the related species *X. tropicalis* provides additional evidence for the instability of these sequences. In fact, three out of the six repetitive DNAs are absent from the DNA of *X. tropicalis*, which represents the most primitive type of *Xenopus*[18]. All these findings suggest that the repetitive sequences within the *A1* genes are unstable elements.

The finding that the repetitive sequences in the oestrogen-inducible *A1* gene are also present in transcriptional units active in the absence of oestrogen (Fig. 4) argues against a simple model in which the six intron-specific repetitive sequences are regulatory elements involved in oestrogen-inducible gene expressions. Future experiments may reveal that repetitive sequences are common in many transcribed introns. In fact, repetitive sequences have recently been detected within introns of the conalbumin gene[11] and the ovalbumin-like *X* gene[19] of the chicken. Because different cells express specific classes of genes, a widespread occurrence of repetitive DNA within transcriptional units would explain the well-documented fact that different tissues contain characteristic sets of repetitive DNA transcripts (reviewed in ref. 1).

All our data on the repetitive DNA within the *A1* gene imply that these sequences reside in this gene without having any obvious function. They might represent some kind of selfish or parasitic DNA[20,21]. However, until further analyses are performed, it remains possible that these repetitive sequences or at least some of them, are essential for gene regulation according to the model proposed by Britten and Davidson[1].

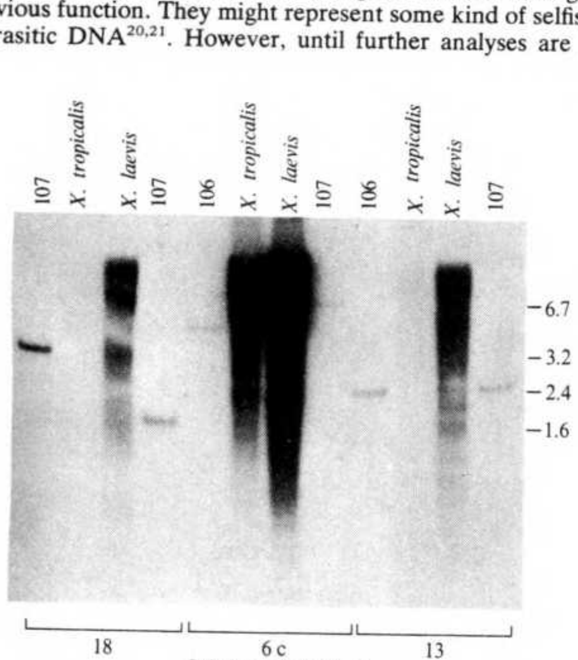


Fig. 3 Distribution of the intron-specific repetitive sequences throughout the genome. Aliquots (10 μg) of *X. laevis* and *X. tropicalis* DNA were restricted with *Eco*RI, separated on a 1% agarose gel, transferred to nitrocellulose filters and hybridized with ^{32}P-labelled inserts of the subclones 6c, 13 and 18. The inserts were cut out by *Pst*I digestion and isolated electrophoretically as described elsewhere[17]. Hybridization was as in Fig. 2. Genomic clones λXlv106 and λXlv107[5] were included as markers in single copy amounts and the molecular weight of these markers are given in kilobases.

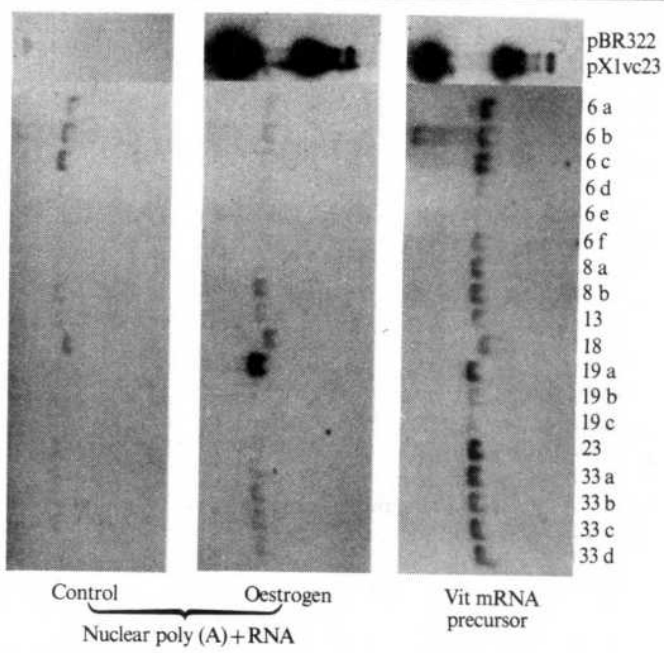


Fig. 4 Transcripts of intron-specific repetitive DNA sequences. Similar blots with intron subclones as in Fig. 2 were hybridized with nuclear polyadenylated liver RNA enriched for vitellogenin mRNA precursor[13], nuclear polyadenylated liver RNA from an oestrogen-treated female[13] and nuclear polyadenylated liver RNA from control males not stimulated with oestrogen. The RNAs were labelled *in vitro* to a specific radioactivity of $2\text{–}4\times10^7$ c.p.m. per μg as follows: 2 μg RNA was partially hydrolysed for 30 min in 0.2 M NaOH at 0 °C (ref. 24), neutralized with HCl, adjusted in a final volume of 50 μl to 120 mM Tris-HCl (*p*H 7.6), 10 mM $MgCl_2$ and 5 mM dithiothreitol, incubated with 50 μCi [γ-^{32}P]ATP (Amersham, PB 10168) and 5 units T_4 polynucleotide kinase (Boehringer) for 30 min at 37 °C and purified on a small Sephadex G-50 column. Hybridization was as in Fig. 2. If the RNA was hydrolysed in control experiments with 0.5 M NaOH before hybridization, no positive reactions were observed, excluding the possibility that the hybridization is due to contaminating DNA in our RNA preparation.

We thank Professor M. Fischberg and the 'Centre Suisse de Recherches Scientifiques en Côte d'Ivoire', Adiopodoumé, for supply of *Xenopus tropicalis*, and Drs I. B. Dawid, R. Jaggi and B. Westley for critical discussions and comments on the manuscript. This work was supported by the Swiss National Foundation and the 'Stiftung zur Förderung der wissenschaftlichen Forschung an der Universität Bern'.

Received 24 November 1980; accepted 5 March 1981.

1. Davidson, E. H. & Britten, R. J. *Science* **204,** 1052–1059 (1979).
2. Fedoroff, N., Wellauer, P. K. & Wall, R. *Cell* **10,** 597–610 (1977).
3. Scheller, R. H., Constantini, F. D., Kozlowski, M. R., Britten, R. J. & Davidson, E. H. *Cell* **15,** 189–203 (1978).
4. Constantini, F. D., Britten, R. J. & Davidson, E. H. *Nature* **287,** 111–117 (1980).
5. Wahli, W., Dawid, I. B., Wyler, T., Weber, R. & Ryffel, G. U. *Cell* **20,** 107–117 (1980).
6. Wahli, W. & Dawid, I. B. *Proc. natn. Acad. Sci. U.S.A.* **77,** 1437–1441 (1980).
7. Tata, J. R. & Smith, D. F. *Recent Prog. Horm. Res.* **35,** 47–95 (1979).
8. Shapiro, D. J. & Baker, H. J. in *Ontogeny of Receptors and Reproductive Hormone Action* (eds Hamilton, T. H., Clark, J. H. & Sadler, W. A.) 309–330 (Raven, New York, 1979).
9. Wahli, W., Dawid, I. B., Ryffel, G. U. & Weber, R. *Science* (in the press).
10. Fritsch, E. F., Lawn, R. M. & Maniatis, T. *Cell* **19,** 959–972 (1980).
11. Cochet, M. *et al. Nature* **288,** 567–574 (1979).
12. Davidson, E. H. & Britten, R. J. *Q. Rev. Biol.* **48,** 565–609 (1973).
13. Ryffel, G. U., Wyler, T., Muellener, D. B. & Weber, R. *Cell* **19,** 53–61 (1980).
14. Ryffel, G. U., Wahli, W. & Weber, R. *Cell* **11,** 213–221 (1977).
15. Baker, H. J. & Shapiro, D. J. *J. biol. Chem.* **252,** 8428–8434 (1977).
16. Wahli, W. *et al. Cell* **16,** 535–549 (1979).
17. Felber, B. K. *et al. Eur. J. Biochem.* **105,** 17–24 (1980).
18. Bisbee, C. A., Baker, M. A., Wilson, A. C., Hadzi-Azimi, J. & Fischberg, M. *Science* **195,** 785–787 (1977).
19. Heilig, R., Perrin, F., Gannon, F., Mandel, J. L. & Chambon, P. *Cell* **20,** 625–637 (1980).
20. Doolittle, W. F. & Sapienza, C. *Nature* **284,** 601–603 (1980).
21. Orgel, L. E. & Crick, F. H. C. *Nature* **284,** 604–607 (1980).
22. Wieringa, B. *et al. Nucleic Acids Res.* **7,** 2147–2163 (1979).
23. Grunstein, M. & Hogness, D. S. *Proc. natn. Acad. Sci. U.S.A.* **72,** 3961–3965 (1975).
24. Wahli, W., Wyler, T., Weber, R. & Ryffel, G. U. *Eur. J. Biochem.* **86,** 225–234 (1978).

Isolation of virus from brain after immunosuppression of mice with latent herpes simplex

Lorne Kastrukoff*, Carol Long, Peter C. Doherty, Zofia Wroblewska & Hilary Koprowski

The Multiple Sclerosis Research Center at The Wistar Institute of Anatomy and Biology, and the Department of Neurology, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

Herpes simplex virus (HSV) is usually present in a latent form in the trigeminal ganglion of man[1-3]. Various stress factors may induce virus reactivation, which is manifest by a lip lesion (innervated from the trigeminal ganglion) and the production of infectious virus. The considerable experimental efforts to define the conditions that lead to the reactivation of latent HSV have concentrated on isolating virus either from the original extraneural site of virus inoculation, or from cell-free homogenates of sensory ganglia from latently infected animals[4-15]. Recent DNA hybridization experiments resulted in the demonstration of the presence of HSV genomes in the brain tissue of both latently infected mice, and of humans who showed no clinical symptoms of HSV (ref. 16 and N. Fraser, personal communication). This led us to consider the possibility that HSV may be present in brain tissue as the result of either reactivation of the virus in brain cells or the passage of reactivated virus from trigeminal ganglia through the brain stem to the brain. The presence of infectious HSV in brain tissue has not previously been demonstrated; yet this could be a factor in chronic, relapsing neurological diseases such as multiple sclerosis. We have now shown experimentally that mice carrying latent HSV in their trigeminal ganglia may, following massive immunosuppression, express infectious virus in the central nervous system (CNS).

Male C3H mice 10–12 weeks of age were lip-inoculated with HSV type 1 (HSV-1) (1×10^7 plaque-forming units (PFU) of laboratory strain 2) as previously described[10]. Four weeks later, surviving animals were divided into two groups (1 and 2) of 20 animals each. Groups 1 and 3 (the latter an uninfected control group of 20 animals) were thymectomized, and group 2 was sham thymectomized. All three groups were irradiated (850 rad) 8 weeks later, and reconstituted with $1-2 \times 10^7$ C3H bone marrow cells that had been treated with anti-Thy-1.2 antibody and complement ('B' mice). Groups 1 and 3 also received intraperitoneally (i.p.) 0.5 ml of a 1:10 dilution of antithymocyte serum (ATS) on days 4, 6 and 8 post-irradiation. Animals in this series were followed clinically, and serial serum samples were obtained for analysis of anti-HSV antibodies measured by solid phase radioimmunoassay (RIA), with HSV-infected CV_1 (indicator) cells as the immunoabsorbent.

The RIA was done as described elsewhere[17]. Briefly, appropriate dilutions of antisera were incubated with immunoadsorbent (in 96-well flat-bottomed tissue culture plates in which 26-h post-infection confluent monolayers of CV_1 cells were fixed with 0.15% glutaraldehyde). After 90 min of incubation at room temperature, the plates were washed and ^{125}I-labelled rabbit anti-mouse $F(ab')_2$ antibody was added. After the final incubation and washes, each well was solubilized in detergent, and the amount of antibody bound determined by γ counting.

Animals in group 1 became lethargic and ataxic between days 5 and 14 post-irradiation whereas those in groups 2 and 3 remained healthy. However, all mice in group 1 recovered and when the study was terminated on day 57, their mortality rate was the same as that for group 3. The animals were bled from the retro-orbital sinus at various times after irradiation and their sera fractionated for determination by RIA of the presence of anti-HSV antibodies in various clones of immunoglobulins. Results of serial anti-HSV antibody determinations are shown in Fig. 1. The major difference between groups 1 and 2 was the level of anti-HSV antibody circulating at the end of the study. Antibody was detected in sera from group 2 but not from group 1 animals on day 57.

Fragments of trigeminal ganglia, lip mucosa and brain tissue were minced and co-cultivated with CV_1 cells. If a cytopathic effect (CPE) was observed in the culture of indicator cells, the virus was passaged and identified in a neutralization test by inhibition of the CPE by anti-HSV immune rabbit serum. At the end of the experiment (day 57), virus was isolated from the trigeminal ganglia of 91% of group 1 and 93% of group 2 survivors but not from the trigeminal ganglia of group 3 mice. Virus was not recovered from the brains of mice from any group.

The next step was to investigate whether HSV could be isolated from the brain tissue of animals when they became acutely ill. We therefore infected a second series of mice (30 per group) with HSV and (after irradiation and immunosuppression) treated them as in the first experiment, the only difference being that animals were killed for sampling on selected days after irradiation (Tables 1, 2). HSV was isolated from the trigeminal ganglia of most of the mice that had been latently infected with HSV (groups 1 and 2, Table 1). Virus was present in the brain tissue in 50% of the mice from group 1 sampled on days 4 and 7 (Table 1), but not in the brain tissue of any other animals. Recrudescence of virus in the CNS thus seems to be dependent on massive immunosuppression, and is accompanied by the development of transient, but not fatal, neurological symptoms. It could be argued that reactivation of HSV might have occurred in tissue culture rather than *in vivo*; however, this is only remotely possible as on many occasions cultivation of CNS tissue containing the viral genome did not result in the isolation of infectious HSV (N. Fraser, personal communication).

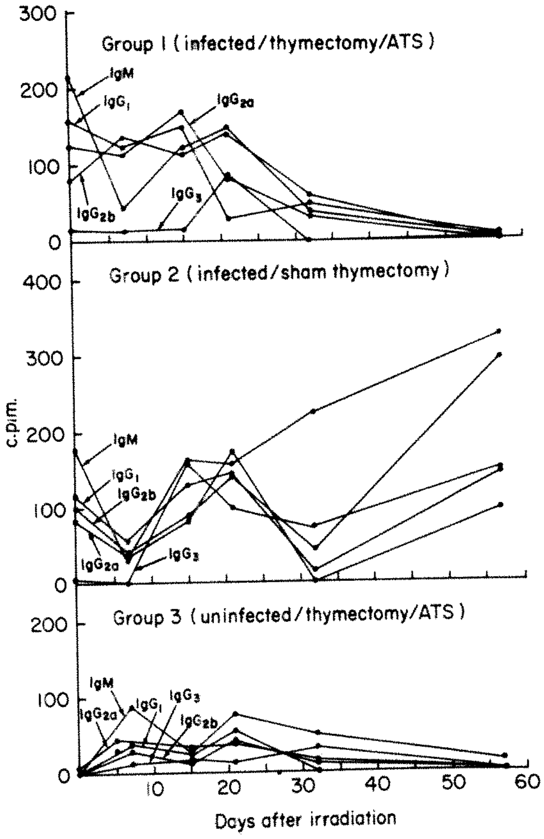


Fig. 1 Levels of serum antibody in latently infected and normal mice that were irradiated (850 rad) and reconstituted with bone marrow 8 weeks after thymectomy or sham thymectomy. Mice in groups 1 and 2 were also dosed with antithymocyte serum (ATS) 4, 6 and 8 days after irradiation.

---

* Present address: The Division of Neurology of the Department of Medicine, University of British Columbia, Vancouver, British Columbia, Canada V6T 1W5.

0028-0836/81/220432—02$01.00

Table 1 Recovery of virus from the brains of mice latently infected with HSV after severe immunosuppression

	Treatment						Virus isolation*			
Group	Latent HSV	ATx	850 rad	ATS	Days after 850 rad	No. of mice sampled	Lip	Trigeminal ganglia	Brain	Negative
1	+	+	+	+	4	8	0	8	4	0
	+	+	+	+	7	8	1	8	5	0
	+	+	+	+	10	7	1	4	0	2
2	+	−	+	−	4	10	0	9	0	1
	+	−	+	−	7	10	1	8	0	1
	+	−	+	−	10	9	2	6	0	1
3	−	+	+	+	4	10	0	0	0	10
	−	+	+	+	7	10	0	0	0	10
	−	+	+	+	10	10	0	0	0	10

Male C3H mice were inoculated in the lip with 1.0×10^7 PFU of HSV (groups 1 and 2), thymectomized (ATx) 4 weeks later (groups 1 and 3), irradiated and reconstituted with bone marrow after a further 8 weeks (all groups) and injected with ATS 4, 6 and 8 days later.

* Tissues were minced and virus isolated by co-cultivation with CV_1 cells. The identity of the brain isolates was determined by neutralization. In all cases, if virus was isolated from the lip or brain, it was also isolated from the trigeminal ganglia.

Table 2 Persistence of antibody in thymectomized, 850-rad HSV-infected mice

		Level of serum antibody				
Group	Day after irradiation	IgM	IgG1	IgG2a	IgG2b	IgG3
1	4	228±107	436±125	417±43	407±173	162±58
	7	192± 60	389±135	401± 4	284± 49	121±18
	10	398	403	263	166	258

The results are expressed, minus background, as mean c.p.m.±s.d. The mice are those from group 1, Table 1.

The effect described above presumably reflects the removal of both 850 rad-sensitive and radiation-resistant (ATS-sensitive) memory T cells. Serum antibody levels were relatively high in the immunosuppressed mice when virus was isolated from the CNS (Tables 1, 2). The possibility thus exists that conditions (such as stress or intercurrent virus infections) that might cause a transient specific defect in HSV-specific cell-mediated immunity[18] could result in the emergence of infectious virus in the CNS. This could then be controlled by non-T cell-dependent mechanisms; virus was isolated from the brains of mice on day 7, but not on day 10 after immunosuppression. New T cells would not be expected to emerge from the thymus for 2–3 weeks. Furthermore, persistence of specific circulatory antibody of any class (Fig. 1) does not seem to be essential for preventing virus reactivation.

Received 12 November 1980; accepted 19 March 1981.

1. Bastian, F. O., Rabson, A. S., Yee, C. L. & Tralka, T. S. *Science* **178,** 306–307 (1972).
2. Baringer, J. R. & Swoveland, P. *New Engl. J. Med.* **288,** 648–650 (1973).
3. Warren, K. G. *et al. Lancet* **ii,** 637–639 (1977).
4. Anderson, W. A., Margruder, B. & Kilbourne, E. *Proc. Soc. exp. Biol. Med.* **107,** 628–632 (1961).
5. Underwood, G. & Weed, S. *Infect. Immun.* **10,** 471–474 (1974).
6. Blyth, W., Hill, J., Field, H. & Harbour, D. *J. gen. Virol.* **33,** 547–550 (1976).
7. Hurd, J. & Robinson, R. *J. Antimicrob. Chemother.* **3,** 99–106 (1977).
8. Hill, T., Blyth, W. & Harbour, D. *J. gen. Virol.* **39,** 21–28 (1978).
9. Walz, M., Price, R. & Notkins, A. *Science* **184,** 1185–1187 (1974).
10. Stevens, J., Cook, M. & Jordan, M. *Infect. Immun.* **11,** 635–639 (1975).
11. Openshaw, H., Asher, L., Wohlenberg, C., Sekizawa, T. & Notkins, A. *J. gen. Virol.* **44,** 205–215 (1979).
12. Hill, T., Field, H. & Blyth, W. *J. gen. Virol.* **28,** 341–353 (1975).
13. Nesburn, A., Green, M., Radnoti, M. & Walker, B. *Infect. Immun.* **15,** 772–775 (1977).
14. Price, R. & Schmitz, J. *Infect. Immun.* **19,** 523–532 (1978).
15. Kurata, T., Kurata, K. & Aoyama, Y. *Jap. J. exp. Med.* **48,** 427–435 (1978).
16. Sequiera, L. *et al. Lancet* **ii,** 609–612 (1979).
17. Frankel, M. E. & Gerhard, W. *Molec. Immun.* **16,** 101–106 (1979).
18. Howes, E., Taylor, W., Mitchison, N. & Simpson, E. *Nature* **277,** 67–68 (1979).

Disorientation of inexperienced young pigeons after transportation in total darkness

Wolfgang Wiltschko & Roswitha Wiltschko

Fachbereich Biologie der Universität, Zoologie, Siesmayerstraße 70,
D 6000 Frankfurt aM, FRG

Leask's proposal[1] that magnetic field detection in birds is accomplished by an optical double-resonance process in the rhodopsin molecules in the retina, a spatially well ordered cellular array, is intriguing. Being an axial rather than a polar phenomenon, it is consistent with observations that birds do not make use of the polarity of the magnetic field[2]. The hypothesis nevertheless requires that light is necessary for sensing magnetic fields, which has prompted us to perform the homing experiments with pigeons reported here. It is known that transportation in a distorted magnetic field causes an increase in scatter, frequently even random, in the initial orientation of young pigeons[3–6], and we sought to discover whether transportation in total darkness would have a similar effect. The results are consistent with Leask's hypothesis but their explanation is not unambiguous. Plainly, however, transportation in total darkness, like the disruption of magnetic or olfactory[3,7] information en route, is another way of preventing the collection of meaningful orientation information during the outward journey to the release site.

Because the effect of treatments during displacement is most pronounced in tests with very young birds without any previous training experience[5,6], we used only untrained birds younger than 12 weeks in the tests reported here. We chose a release site 65 km SSW of the home loft (home direction 17°). Depending on local traffic, the driving time was 50–75 min. On sunny days, the birds were released singly, alternating experimentals and controls. They were observed with 10 × 40 binoculars until they vanished from sight; their departure bearings were recorded to the nearest 5° and later statistically evaluated as described by Batschelet[8]. Homing times and homing success could not be evaluated because the distance of 65 km proved too large for our young birds, and very few returned.

Table 1 Effect of transportation in total darkness on orientation of pigeons after release

	Controls			Experimentals		
Test no.	*n*	α (deg)	*r*	*n*	α (deg)	*r*
1	9^s	6	0.81*	9	286	0.20‡
2	9^s	345	0.70†	10	272	0.14‡
3	9^l	11	0.76†	8	41	0.23‡
4	10^l	345	0.82*	10	138	0.29‡

With the controls, s indicates birds that sat in the dark for 30 min after arriving at the release site, and l indicates birds transported in the lighted box (see text). Significance by the Rayleigh test: *, $P < 0.001$; †, $P < 0.01$; ‡, $P > 0.05$, not significant.

028-0836/81/220433—02$01.00

To demonstrate the effect of transportation in an altered magnetic field, the birds were transported in wooden baskets, the walls and roof of which consisted of small wooden poles allowing the birds to see the outside. For the controls this basket was placed on the floor of the transportation vehicle, a Volkswagen bus with windows all round, where the magnetic field was only very slightly distorted ($<10\%$) by the metal frame of the car. The experimentals were transported in a second vehicle of the same type, but were placed in the centre of a pair of Helmholtz coils which produced a horizontal magnetic field of ~ 0.4 G pointing towards the front of the car. This field added to the local field as we described before[5,6], resulting in a strongly distorted total field whose characteristics depended on the direction of driving. After arriving at the release site, both baskets were placed on the ground, giving the birds access to the natural environmental stimuli. In four such tests the normally transported control birds were significantly homeward oriented, whereas the birds transported in the artificially altered magnetic field showed no significant directional tendency ($P<0.01$, Mardia Watson Wheeler test with standardized mean directions). The results are summarized in Fig. 1.

To test whether transportation in total darkness produced a similar effect, we placed the basket of experimental birds in a light-tight wooden box, which was transported on the rooftop baggage rack of the vehicle and ventilated through shielded air slits by the natural air stream. Thus, the magnetic field was undisturbed and the birds could breathe the natural outside air during the outward journey (a condition related to findings of Papi *et al.*[3,7]). After arriving at the release site, the basket with the experimental birds was taken out of the dark box and placed on the ground. We used two types of control: in releases 1 and 2, the control birds were transported in the wooden basket like the control birds in the first experiment, but were placed in the dark box for 30 min after arrival at the release site (Fig. 2*a*, open symbols). In releases 3 and 4, the controls were also transported in a similar daylight-tight wooden box on the roof of the vehicle but two small battery-operated light bulbs provided a dim light (3.4 lx) inside the box (Fig. 2*a*, solid symbols). The data for all four tests are given in Table 1. The control birds were significantly oriented, whereas the birds transported in total darkness showed random behaviour. These data are summarized in Fig. 2; the difference in scatter is significant ($P<0.01$).

The control results indicate that this disorientation was not caused by simply staying in total darkness for a certain amount of time, nor by being transported visually isolated from the outside. Transportation in total darkness, as in a distorted magnetic field, seemed to make the gathering of necessary information during the outward journey impossible. (In release 4 we controlled against psychological effects by habituating the designated experimental birds to being driven around in the dark box before the critical test, but this did not seem to help their orientation.)

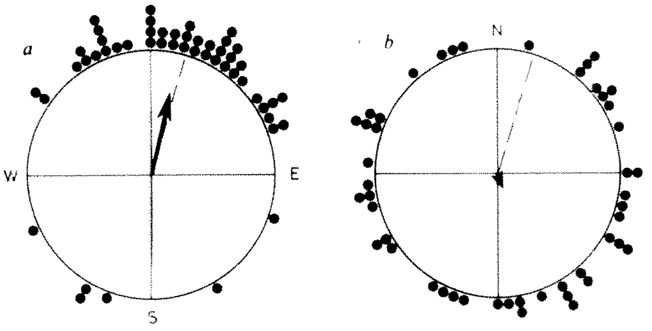


Fig. 1 Summary of the four tests demonstrating the effect of a distorted magnetic field during the outward journey. *a*, The normally transported controls were significantly oriented: $n=50$, $\alpha=13°$, $r=0.66$ ($P<0.001$, Rayleigh test). *b*, The birds transported in the experimentally altered magnetic field did not show any directional preference: $n=48$, $\alpha=156°$, $r=0.13$ ($P\gg0.05$). The departure bearings are represented by the symbols at the periphery of the circles; the mean vectors are given as arrows with the length r drawn proportional to the radius of the circle $=1$. The home direction, 17° NNE, is marked by a dashed line.

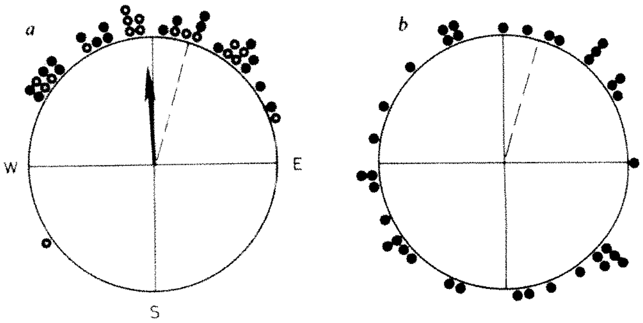


Fig. 2 Summary of the four tests demonstrating the effect of total darkness during the outward journey. *a*, Both types of control (see text) were significantly oriented: $n=37$, $\alpha=357°$, $r=0.76$ ($P<0.001$). *b*, The birds transported in total darkness showed random behaviour: $n=37$, $\alpha=137°$, $r=0.01$ ($P\gg0.05$). Symbols as in Fig. 1.

Our findings are consistent with a major prediction derived from the optical resonance hypothesis. But, as 'negative results', they are open to various explanations, since we do not know the exact cause of the disoriented behaviour of our experimental birds. 'Orientation' as recorded in homing experiments with birds is a highly complex behaviour, the observed disorientation shows that one of the functions leading to the behavioural output had been disrupted, but we cannot tell at what level, nor can we discriminate between an effect on the magnetic receptor and other effects (like, for example, the birds falling asleep or feeling sick when transported in darkness). This dilemma must be considered when our findings are discussed in the light of Leask's[1] hypothesis, the more, since other orientation experiments with caged migratory birds have yielded ambiguous results (unpublished).

Had the outcome of our experiments been reversed, that is, had transportation in total darkness not affected the birds' orientation, it would have clearly shown that Leask's[1] model is incorrect, at least for magnetoreception in birds. The results, as they turned out, do not allow a definite conclusion about the model: they suggest that it might apply, but other specific or nonspecific interferences with the orientation behaviour are equally possible.

Nonetheless, transportation in total darkness is a methodically simple way of preventing pigeons from obtaining orientation information during the outward journey, and this will help any future study of the significance of route-specific information compared with site-specific information in the orientation system of birds[9].

This work was supported by the Deutsche Forschungsgemeinschaft, all calculations were carried out by the Hochschulrechenzentrum of the Universität Frankfurt aM. We thank our helpers, especially E. Franke, H. Golle, M. Grüter and V. Reichard, for assistance with the releases, and M. J. M. Leask for critically reading an early draft of this manuscript.

Received 5 January; accepted 25 March 1981.

1. Leask, M. J. M. *Nature* **267**, 144–145 (1977).
2. Wiltschko, W. & Wiltschko, R. *Science* **176**, 62–64 (1972).
3. Papi, F., Ioalè, P., Fiaschi, V., Benvenuti, S. & Baldaccini, N. E. *Animal Migration, Navigation and Homing* (eds Schmidt-Koenig, K. & Keeton, W. T.) 65–77 (Springer, Berlin, 1978).
4. Kiepenheuer, J. *Naturwissenschaften* **65**, 113 (1978).
5. Wiltschko, R. & Wiltschko, W. *Naturwissenschaften* **65**, 112–113 (1978).
6. Wiltschko, R., Wiltschko, W. & Keeton, W. T. *Animal Migration, Navigation and Homing* (eds Schmidt-Koenig, K. & Keeton, W. T.) 152–162 (Springer, Berlin, 1978).
7. Papi, F. *Verh. dt. zool. Ges.* **1976**, 184–205 (1976).
8. Batschelet, E. *Statistical Methods for the Analysis of Problems in Animal Orientation and Certain Biological Rhythms* (AIBS, Washington DC, 1965); *Animal Orientation and Navigation* (eds Galler, S.R. *et al.*) 61–63 (NASA SP-262, Washington DC, 1972).
9. Wiltschko, W. & Wiltschko, R *Oikos* **30**, 177–187 (1978).

BOOK REVIEWS

Meeting of physics and metaphysics

Abner Shimony

THIS book is David Bohm's attempt to effect a meeting of East and West. In the basic principles of his metaphysics and epistemology, the East predominates. The world is an undivided whole, and any analysis of it into parts (things, atoms, sensations) is an erroneous fragmentation. Unlike the static monism of Parmenides, the monism of Bohm emphasizes process: "all *is* flux" (p.48). A primary instance of violating the wholeness of the world is the dichotomy of mind and matter. It is incorrect to regard knowledge as a correspondence between thoughts and things, because both of these termini of the hypothetical correspondence relation are abstracted from the total process. Bohm proposes instead "the notion that thought is a sort of 'dance of the mind' which functions indicatively" (p.55). Ordinary language with its characteristic subject–verb–object structure is inadequate both to the holistic and the process character of reality. Consequently, Bohm proposes a new mode of discourse called "the rheomode", from the Greek word meaning "to flow", in which the verb rather than the noun is given a primary role.

Bohm acknowledges, however, that the West has something to contribute to wisdom, since "measure is insight into a secondary and dependent but nevertheless necessary aspect of reality" (p.23). Indeed, if it were not for his effort to relate modern physics to his metaphysics there would be little reason to single this book out for attention from among numerous variant versions of holistic metaphysics. One might be content with the succinct exposition of Enoch Soames: "Life is web, and therein nor warp nor woof is, but web only" (in Max Beerbohm's *Seven Men;* Heinemann, 1919).

Over three decades Bohm has made distinguished contributions to the foundations of quantum mechanics. In the present book his position regarding quantum mechanics is a complicated synthesis of his previous ideas. He continues to search for a theory of hidden variables on a sub-quantum level. He recognizes, however, that in situations of the type considered by Einstein, Podolsky and Rosen, in which a non-factorizable wave function is used to describe a pair of spatially separated particles, no *local* hidden variables theory is capable of agreeing with all the experimental predictions of quantum mechanics. (The argument which he gives to this effect on pp.72–73 is unfortunately very elliptical and could not be followed by any reader who is not acquainted with the work on local hidden variables theories of J.S. Bell and his followers.) Bohm is willing, however, to accept non-local hidden variables theories. He thus seems to believe that some of the features of present-day quantum theory — especially holism and non-locality — will remain embedded permanently in the *Weltanschauung* of physics even after the development of a sub-quantum physics; indeed, two of his seven chapters contain "Quantum theory as an indication of a new order in physics" as part of their titles. Clearly, the features of quantum mechanics which he expects to be perpetuated are precisely those which mesh with his metaphysics of holism and flux.

Wholeness and the Implicate Order. By David Bohm. Pp.224. ISBN 0-7100-0366-8. (Routledge & Kegan Paul: 1980.) £8.95, $25.

Since the causal structure of relativistic space-time appears to be violated in any non-local hidden variables theory, the following question is posed: why is it not possible to transmit signals faster than light? Bohm answers that only random motions on a sub-quantum level are superluminal, whereas the average motions of particles, which are relevant in signal propagation, are governed by relativity theory. This answer can be assessed only when the physical implications are worked out in detail and appropriate experiments are performed.

What is most puzzling about Bohm's philosophy is the problem which it shares with all varieties of holism: how does an undivided existence manage to exhibit itself in fairly well articulated parts and aspects? Bohm attempts to solve this problem by the interplay of the "implicate order" and the "explicate order" (Chapters 6 and 7) In the implicate order "the totality of existence is enfolded within each region of space and time" (p.172).

> Nevertheless, . . . the overall law (holonymy) may be assumed to be such that in a certain sub-order, within the whole set of implicate order, there is totality of forms that have an approximate kind of recurrence, stability and separability . . . The special distinguished sub-order indicated above, which is the basis of the possibility of this manifest world, is then, in effect, what is meant by the explicate order [p.186].

Bohm's favourite illustration is the hologram, in which a whole illuminated structure is mapped by Fourier optics onto every finite region of a photographic plate, with the result that the original optical display can be reconstructed (though with some loss of detail) even if all but a small part of the photographic plate is destroyed. The optical display reconstructed by interfering laser beams is the "explicate order", which makes manifest the "implicate order" in the photographic plate. Of course, this illustration is little more than a metaphor, since holography can be explained in terms of classical electromagnetic theory, which does not represent reality as an undivided whole. (Each component of the electromagnetic field at each point on a space-like surface is a separate degree of freedom, and all of these infinitely-many degrees of freedom can in principle be assigned independent values.) Nevertheless, the metaphor of the hologram is suggestive, and Bohm's conception of the implicate and explicate orders does indeed seem to provide him with an inspiring heuristic for revising physical theory.

Bohm's programme is ambitious, and in the long run it must be assessed by concrete achievements. He hopes to remove the mystery from the reduction of wave packets; he wishes to eliminate the divergences which bedevil relativistic quantum field theory and to dispense with techniques such as renormalization. Since the quantization of the gravitational field suggests that it may be meaningless to speak of spatial distances less than 10^{-33}cm, he hopes to be able to throw new light upon the structure of space-time in the small. And he expects new phenomena to be observed when the "immense background of energy" in so-called empty space is properly probed. The sketch which Bohm provides of the mathematical apparatus of his programme on pp.155–171, is, unfortunately, so cryptic that I can make no judgements about its prospects. I can only say that he has achieved enough as a physicist to deserve both time for pursuing his programme and attention for his claims of concrete results.

It is appropriate to compare Bohm's book with Fritjof Capra's *The Tao of Physics* (Wildwood House, 1975), which also advocates the holistic metaphysics of the East and links it with modern physics. Capra, incidentally, quotes approvingly from an important paper of Bohm and Basil Hiley. Capra's book is pleasant to read, for he writes fluently, quotes charming *haikus* and Zen *koans*, and even offers some beautiful photographs of sculpture and calligraphy. Bohm's style, by contrast, is ponderous and his exposition abstract. What I find unattractive in Capra's book is the way in which he draws up a brief that "the principal theories and models of modern physics lead to a view of the world which is internally consistent and in perfect harmony with the views of Eastern mysticism". He relies uncritically upon the controversial bootstrap theory of strong interactions, because it fits his metaphysics; he dissolves the precise meaning of technical terms in order to establish analogies; and he plays down discrepancies between points of view when they threaten his parallels. Bohm's book has some of the same tendencies, but to a much lesser degree. *Wholeness and the Implicate Order* conveys a sense of work in progress, which aims at a distantly glimpsed ideal of the unification of all the aspects of the world, and it is refreshingly free from claims that the ideal has already been achieved. The feeling of struggle in Bohm's book is its most appealing feature.

Abner Shimony is Professor in the Departments of Philosophy and Physics at Boston University.

IRCS JOURNALS

Keeping pace with science

IRCS Medical Science
St. Leonard's House
St. Leonardgate,
Lancaster, U.K.
A member of Elsevier Science Publishers

Circle No. 14 on Reader Enquiry Card.

Birthday party for the skinned fibre

C.R. Bagshaw

Muscle Contraction: Its Regulatory Mechanisms. Edited by S. Ebashi *et al.* Pp.549. ISBN 3-540-10411-9/0-387-10411-9. (Japan Scientific Societies Press, Tokyo/Springer-Verlag: 1980.) DM108, $63.80.

DURING October 1979 an international symposium on muscle contraction was held near Mount Fuji in honour of two eminent Japanese scientists, Professor Hiroshi Kumagi and Professor Reiji Natori. In 1955 Kumagi and colleagues identified the "relaxing factor" of muscle as a microsomal ATPase — the now-familiar Ca^{2+} pump of the sarcoplasmic reticulum. Since then he has been a central figure in organizing and stimulating muscle research in Japan. Natori's "skinned fibre" needs little introduction, having provided a unique preparation amenable to combined physiological and biochemical manipulations. These symposium proceedings have been published in commemoration of Professor Kumagi's seventy-fifth and the skinned fibre's formal twenty-fifth birthday.

The majority of the 42 contributions comprise summaries of recent work, but without unnecessary duplication of published data. Instead, many authors have introduced the historical events and intuitions which have guided the direction of their research. A.F. Huxley starts the proceedings with a review that makes the library stacks as enlightening as the current journal shelves. He describes the initial reluctance to accept that relaxing agents work by chelating Ca^{2+} — I wonder if we have become too complacent in our use of EGTA. Natori relates how his skinned fibre helped in the evaluation of the excitation–contraction mechanism, but not content to dwell on history he goes on to describe the contribution of connectin to passive elasticity.

The remaining papers range from studies of cross-bridge movement to the cytoskeleton, but predominantly concern regulatory mechanisms. Unfortunately, while the controversies regarding smooth muscle control are outlined, insufficient data are presented to allow a rational assessment. Hartshorne and Adelstein focus on the properties of kinases, while Nonomura's leiotonin data are limited. Watanabe presents the noteworthy result that desensitized molluscan actomyosin does not superprecipitate, yet Szent-Györgyi shows that desensitized fibres contract regardless of Ca^{2+}. Over the past decade superprecipitation has fallen out of favour as a quantitative probe; clearly, caution is required even in its qualitative applications, particularly in preparations which may possess atypical solubility at low ionic strength.

This volume has achieved its aim of bringing together current research on regulatory mechanisms. The readable styles adopted by most authors serve to complement the existing literature, and the editors are to be congratulated on compiling a fitting and useful tribute. □

C.R. Bagshaw is a Lecturer in Biochemistry at the University of Leicester.

Polymers in the lab

Herman Mark

Experimental methods in Polymer Chemistry. By J. F. Rabek. Pp.861. ISBN 0-4712-7604-9. (Wiley: 1980.) £52.80, $145.20.

THIS presentation of a large and important field of modern chemistry is truly remarkable, particularly since it has been compiled and written by a single author.

Two chapters are essentially descriptive, as they ought to be — the overviews of the structure (Chapter 1) and of the morphology (Chapter 26) of polymers; both contain well-selected and educationally impressive sketches and photographs of all aspects of their topics. In addition they make excellent reading. The bulk of the book is really a condensed encyclopaedia of all the basic characteristics of synthetic polymers. In each case (for instance light scattering, Chapter 13) there is first a clear yet concise presentation of the theoretical background, then a description of the best existing instruments with adequate instructions for making sample preparations and, finally, the most recent improvements and modifications of the particular method are discussed. This enumeration starts with dissolution and goes over fractionation, osmometry, ultracentrifugation, viscometry, all spectroscopic methods to the scattering of X-rays, electrons and neutrons, to NMR and the various chromatographic techniques.

Deserving of particular mention are the condensed but clear presentations of the principles of every individual technique and the many long and carefully assembled tables on significant results, for example those on solubility parameters, NMR parameters and ESR parameters. A special feature of the book, and at the time proof of the never-ending patience of the author, is the bibliography which takes up almost 150 pages and lists 7,350 entries.

If anybody wants to get — or stay — in close touch with polymer chemistry, he ought to buy this book and carry it with him all the time; there is no better source of information on the subject. □

Herman Mark is Emeritus Dean of the Polytechnic Institute of New York, and editor of the Journal of Polymer Science.

Henry Stommel and physical oceanography since 1945

K.F. Bowden

Evolution of Physical Oceanography. Edited by Bruce A. Warren and Carl Wunsch. Pp.623. ISBN 0-262-23104-2. (MIT Press: 1981.) $37.50, £23.25.

THIS volume is a tribute, on the occasion of his sixtieth birthday, to Henry Stommel who has been a pre-eminent figure in the development of physical oceanography since the Second World War. He is probably best known for showing, in an early paper published in 1948, that the intensification of currents on the western side of an ocean is due to the variation of the Coriolis parameter, which expresses the effect of the Earth's rotation, with latitude. This is demonstrated clearly by the relatively narrow, fast-flowing Gulf Stream in the North Atlantic and the Kuroshio in the North Pacific Ocean, but it is a dynamical feature of deep currents as well.

V. Cullen

Henry Stommel

Stommel has subsequently made remarkable contributions to practically every aspect of physical oceanography, from mixing in estuaries to the world-wide deep circulation of the oceans. His work has been characterized by its physical insight, expressed in terms of the simplest dynamical models, which others have developed in greater mathematical detail. He has also played a leading part in the move to direct field work in oceanography away from the survey of geographical areas to the study of physical processes.

In their preface the editors explain why, instead of the traditional *Festschrift*, this volume takes the form of a series of broad, comprehensive surveys of aspects of oceanography with which Stommel has been concerned. Together they present a most valuable picture of the evolution of physical oceanography over the past 35 years and its present status. The surveys are preceded by short contributions from five former colleagues, giving an appreciation of Henry Stommel and his work.

The main body of the book is comprised of 18 chapters grouped into four parts. The first, on the general ocean circulation, is the longest. It has surveys of the deep circulation of the world ocean by B.A. Warren, on water masses by L.V. Worthington, the mid-depth circulation by J.L. Reid and the Gulf Stream System by N.P. Fofonoff. In his chapter on the dynamics of large-scale ocean circulation, G. Veronis gives a comprehensive account of the vast amount of theoretical work in this field over the past 30 years. Equatorial currents are clearly described in observation and theory by A. Leetmaa, J.P. McCreary and D.W. Moore. The chapter on estuarine and continental-shelf circulation in the Middle Atlantic Bight, by R.C. Beardsley and W.C. Boicourt, is of wider interest than its title might suggest, as the treatments and results are applicable to many other areas.

Part Two deals with physical processes in oceanography. The first chapter, by J.S. Turner, is on small-scale mixing processes and is followed by that of W. Munk on internal waves and small-scale processes. Long waves and ocean tides are reviewed by M.C. Hendershott, while C. Wunsch discusses the low-frequency variability of the sea, a feature which has long been recognized but has only recently come to be studied with a thoroughness approaching that with which meteorologists have treated the atmosphere. In the only contribution of a biological nature, J.H. Steele discusses the interaction of physical, biological and ecological processes in the distribution of organisms. W.V.R. Malkus rounds off this section with a chapter on the amplitude of convection.

The practical oceanographer will welcome Part Three, "Techniques of Investigation", which opens with a good review of instrumentation and the design of experiments by D.J. Baker. W.S. Broecker discusses geochemical tracers and ocean circulation, dealing mainly with the use of radioisotopes, particularly the data from GEOSECS (Geochemical Ocean Sections Studies). The use of laboratory models in the study of ocean circulation is described by A.J. Faller. In the final section on ocean and atmosphere, H. Charnock surveys the topic of air–sea interaction, while J.G. Charney and G.R. Flierl describe ocean analogues of large-scale atmospheric motions, concentrating mainly on transient motions.

In a volume of this kind there is bound to be a certain unevenness among the contributions, although here the treatment is more consistent than in most collective works. All chapters are of high quality and comprehensive in scope, the variations being in depth and in the amount of mathematical detail in the chapters on theoretical topics. This volume has succeeded admirably in its purpose of paying tribute to a great oceanographer and providing a comprehensive and inspiring survey of physical oceanography today. □

K.F. Bowden is Professor of Oceanography at the University of Liverpool.

Surface tension

Peter Schofield

Statistical Mechanics of the Liquid Surface. By Clive A. Croxton. Pp.368. ISBN 0-4712-7663-4. (Wiley: 1980.) £25, $68.75.

THE first thing to be said about this volume is that the title is wrong. The book is mainly concerned with the *molecular theory* of liquid surfaces, that is with methods of calculation of the properties of liquid–vapour interfaces either from the intermolecular forces and molecular distribution functions (the Kirkwood-Buff approach) or, to a lesser extent, from the old Van der Waal's theory and its modern counterpart in terms of the density gradients and density fluctuations (the Yvon–Triezenberg–Zwanzig approach). In both cases, the derivation from statistical mechanics is given in a manner which the uninitiated reader will find difficult to follow.

The fundamental question about surface tension is whether indeed it has an intrinsic value for a given fluid. The width of the interface is determined by the amplitude of thermal capillary wave fluctuations, which is unbounded in the limit of infinite volume and zero gravity since there is then no cost in free energy in a rotation of the whole system. It is on this quicksand that the molecular theory is built!

Nevertheless, considerable success has been achieved in predicting the properties of the liquid-vapour interface by ignoring such problems, and Clive Croxton has done a service in bringing together this review, with a comprehensive bibliography, of calculational methods for several types of fluid. Following two introductory chapters, the subjects include the surfaces of molecular, metallic and quantum liquids, liquid crystals and polymer adsorption at solvent surfaces. The question of molecular orientation in an interface and surface potentials is of considerable practical importance, and in this respect there is a good chapter on water.

As in many fields of physics and physical chemistry, computer simulation methods have a major role to play in bridging the gap between theory and experiment. On the evidence of the chapter devoted to this topic, these techniques were only just approaching the stage of giving definitive results for the liquid surface at the time of writing, though progress has been made since. The situation resembles that for bulk liquids some 20 years ago. There is no doubt that computer simulation can likewise transform our understanding of the molecular properties of surfaces.

Computer simulation, however, will not solve the deeper statistical mechanics problems of co-existing phases of real fluids, such as the conditions for phase separation and critical behaviour. For these, one needs the full, modern apparatus of renormalization group and nucleation theory, but such considerations are outside the scope of this book. □

Peter Schofield is a Group Leader in the Materials Physics Division at Harwell, and has overall responsibility for the Neutron Beam Research Programmes of the Atomic Energy Authority.

The iris as history

J.C. Dean Hart

The Iris in Eighteenth-century Physiology. By Renato G. Mazzolini. Pp.193. ISBN 3-456-81022-9. (Hans Huber, Bern: 1980.) Flexi SwFr. 36, DM 39.

Recent reviews of the historical aspects of iris musculature, notably that of Loewenfeld 1957, when dealing with the eighteenth century use as their first definitive reference the report by Albrecht von Haller, Professor of Anatomy at Göttingen, who in 1753 stated that the iris did not contain muscles. This opinion flatly contradicted the contemporary view that an antagonistic set of muscles was present in the iris, one for constriction of the pupil, the other for dilation. However, no recent author has made mention of the persons or person responsible for initiating the muscular concept or of how it became widely disseminated. Here, in an erudite piece of historical investigation, Mazzolini has tracked down the source of the muscular theory and convincingly demonstrated reasons as to why it became so readily accepted by most authorities on the Continent and as far away as Scotland. He has set out to provide an historical analysis of the statement by Haller that "all muscles are irritable — the iris has no irritability", and has attempted to answer three questions: was this remark original; to what extent did Haller's comments alter views of students on the subject; and how did this change previously held ideas? It would, however, be unfair in this review to reveal all the details of Mazzolini's research, as his is an excellent piece of detection.

Haller began to have serious doubts about the presence of muscles in the iris as early as 1743, since, unlike earlier researchers, he was quite unable to see any such structures. In 1753, his major work was published, *De partibus corporis humani sensilibus et irritabilus*, in which an entirely new concept of physiological thinking was set out based solely on the twin pillars of experimentation and observation. His novel proposal was that the body could be divided into three types of tissue depending on how they reacted to cutting, burning or the application of noxious fluids: those showing irritability (contractability) were muscles; those demonstrating sensibility, nerves; the remainder, having neither of these attributes, were called tela cellularis.

This work proved to be a landmark since it introduced a radical alternative to the earlier scientific approach based on the Galenic tradition of analogy and reason. The case for iris muscles had been argued previously in the form of a syllogism as follows: structures that contain muscles contract; the iris contracts; therefore this tissue contains muscles. Haller, however, believed that analogy was a source of great error and that it was not necessary to devise fabrics which could not be observed. Since irritability of the iris was not detectable in experimental animals when this tissue was cut or after the application of oil of vitreol — and supported by his earlier observation that muscles could not be seen — he felt that the idea of a muscular mechanism of iris movement had to be abandoned. True to the tradition that he had created, Haller proposed instead that the blood vessels of the iris, which could be readily identified using wax injection techniques, became dilated or constricted, thus permitting blood to flow in and out and inducing movements similar to the type known to occur in the erection of the penis. The erectile theory continued to be predominant for the next 70 years or so; only when improved microscopes became available and tissues could be sectioned and suitably stained, in the 1820s, was the sphincter muscle finally identified beyond all reasonable doubt. In the case of the dilator muscle, however, a fierce controversy was only quelled in the early part of the twentieth century when embryological studies revealed that both the dilator and constrictor muscles were derived from neural ectoderm.

In this book, Mazzolini has accomplished the task which he set himself; Haller's hypothesis does appear to have been original and to have radically influenced contemporary views. However, there are a few aspects in the text where further elaboration might have been helpful. Mazzolini has attempted by experimentation and using techniques similar to those employed by eighteenth century anatomists to determine how a number of ideas concerning iris movements were derived, but he has not sought to question the truth of Haller's statement — "the iris is not irritable". This is unfortunate as one gains the impression that Haller's concept remains correct. Modern ophthalmologists would agree that in the Hallerian sense the human iris does in fact exhibit irritability, since if it is accidentally photocoagulated or subjected to excessive manipulation — for example, during cataract surgery — changes in the size of the pupil may occur, albeit quite slowly.

Also, throughout the book, philosophical concepts prevailing during the eighteenth century are discussed which the author believes may have had an influence on attitudes on various workers on reaching their conclusions. But, particularly in the last chapter on *Naturphilosophen*, such theories have not been sufficiently well described for anyone who is not well versed in continental ideologies of the period to achieve an adequate understanding of a number of rather complex propositions. It likewise seems unnecessary to contemplate whether the Scotsman, Robert Whytt (1714–1766) was an unorthodox Stahlian or not, irrespective of Haller's views on this subject, since his reasoning on the nature of the pupillary responses to alterations of illumination is of the very highest quality in his work *An essay on the vital and other involuntary movements of animals,* published in 1751. The fact that he considered the soul or sentient spirit developed an awareness that the retina had become uneasy if too much light impinged upon it and responded by causing pupillary constriction, could be explained equally well on the basis that Whytt, in trying to describe the concept of a reflex, might have been attempting to avoid rousing the animosity of the powerful national religious orthodoxy of the period. Whytt's views were temporarily discredited, not so much because he believed in animism, but because there was insufficient experimental data to support his supposition in terms of Haller's new physiology.

The eighteenth century was a period of profound importance in the development of the medical sciences; it was during this time that the use of experimentation became widely respectable and the mists of archaic speculation began to melt away. Mazzolini is to be congratulated on providing a clear and eminently readable exposition of a period of the history of the iris which has recently suffered almost total neglect. Although speculations relating to the mechanics of the iris might appear to be a somewhat esoteric subject, we are given a clear indication of the wider aspects of physiological investigations that were being pursued during this period, which included the first steps as to how the fundus might be examined *in vivo* and the role of the retina as a light-sensitive tissue. □

J.C. Dean Hart is Head of the Department of Ophthalmology at the University of Bristol.

BOOKS RECEIVED

Mathematics

ADLER, R. J. The Geometry of Random Fields. Pp.280. ISBN 0-471-27844-0. (Wiley: 1981.) £17.50.

BOYCE, W. E. (ed.). Case Studies in Mathematical Modeling. Pp.386. ISBN 0-273-08486-0 (Pitman Publishing: 1981.) £22.

DRAZIN, P. G. and REID, W. H. Hydronamic Stability. Cambridge Monographs on Mechanics and Applied Mathematics. Pp.525. ISBN 0-521-22798-4. (Cambridge University Press: 1981.) £35.

LAPWOOD, E. R. and USAMI, T. Free Oscillations of the Earth. Pp.243. ISBN 0-521-23563-7. (Cambridge University Press: 1981.) £25.

REIN, W. T. Sturmian Theory for Ordinary Differential Equations. Applied Mathematical Sciences, Vol.31. Pp.559. Flexi ISBN 3-540-90542-1. (Springer-Verlag: 1980.) DM 54, $31.90.

Physics

BOUMANS, P. W. J. M. Line Coincidence Tables for Inductively Coupled Plasma Atomic Emission Spectrometry. 2 Vol. set. Pp.951. ISBN 0-08-026243-0-H (Pergamon: 1980.) £104, $250.

BUBE, R. H. Electrons in Solids. An Introductory Survey. Pp.229. ISBN 0-12-138650-3. (Academic: 1980.) $25.

FRAISSARD, J. P. and RESING, H. A. (eds). Magnetic Resonance in Colloid and Interface Science. NATO Advanced Study Institutes Series. C. Mathematical and Physical Sciences, 61. Proceedings of a NATO Advanced Study Institute and the 2nd International Symposium held at Menton, France, June 1979. Pp.710. ISBN 90-277-1153-4. (Reidel: 1980.) Dfl.145, $76.

GARRIDO, L. (ed.). Systems Far from Equilibrium. Lecture Notes in Physics, 132. Sitges Conference on Statistical Mechanics, June 1980, Sitges, Barcelona, spain. Pp.403. Flexi ISBN 3-540-10251-5. (Springer-Verlag: 1980.) DM147, $27.80.

LAWRENCE, J. H. GOFMAN, J. W. and HAYES, T. L. (eds.) Advance in Biological and Medical Physics, Vol. 17. Pp.473. ISBN 0-12-005217-2. (Academic: 1980.) $55.

PAMPLIN, B. R. (ed.). Crystal Growth. 2nd Edn. Pp.609. ISBN 0-08-025043-2. (Pergamon: 1980.) £60, $140.

WOLF, K. B. (ed.). Group Theoretical Methods in Physics. Lecture Notes in Physics, 135. Proceedings of the IX International Colloquium held at Cocoyoc, Mexico, June 1980. Pp.629. Flexi ISBN 3-540-10271-X. (Springer-Verlag: 1980.) DM72 $42.50.

Chemistry

HANNA, L. W. Quantum Mechanics in Chemistry. 3rd Edn. Pp.284. Hbk ISBN 0-8053-3708-3; pbk ISBN 0-8053-3705-9. (Benjamin/Cummings: 1981.) Hbk $14.95; pbk $9.95.

ORCHIN, M. *et al.* The Vocabulary of Organic Chemistry. Pp.609. ISBN 0-471-04491-1. (Wiley: 1981.) £18.75, $43.75.

TEJA, A. S. (ed.). Chemical Engineering and the Environment. Critical Reports of Applied Chemistry, Vol.3. Pp.100. Flexi ISBN 0-632-00693-5. (Blackwell Scientific: 1981.) £8.

Technology

COUNT, B. (ed.). Power from Sea Waves. Based on the Proceedings of a Conference held at University of Edinburgh from 1979. Pp.450. ISBN 0-12-193550-7. (Academic: 1980.) £23.60, $57.

JENKINS, S. H. (ed.). Mediterranean Coastal Pollution. Progress in Water Technology. Proceedings of a Conference held in Palma, Mallorca, September 1979. Pp.870. ISBN 0-08-026058-6. (Pergamon: 1980.) £44.50, $100.

KUO, F. F. Protocols and Techniques for Data Communication Networks. Pp.468. ISBN 0-13-731729-8. (Prentice-Hall: 1981.) £19.45

OLSZEWSKI, M. (ed.). Utilization of Reject Heat. Energy Power and Environment, Vol.9. Pp.208. ISBN 0-8247-1168-8. (Marcel Dekker: 1980.) SwFr.65.

SOCIETE FRANCAISE D'ENERGIE NUCLEAIRE. Colloquium on the Risks of Different Energy Sources. Pp.658. (Gedim, 19 Rue du Grand-Moulin, 42029 Saint-Étienne Cédex: 1980.) Flexi np.

Earth Sciences

GUPTA, H. K. Geothermal Resources: an Energy Alternative. Developments in Economic Geology, 12. Pp.220. ISBN 0-444-41865-2. (Elsevier Scientific: 1980.) $61, Dfl.125.

LYNCH, A. J. *et al.* Mineral and Coal Flatation Circuits. Their Simulation and Control. Developments in Mineral Processing, 3. Pp.291. ISBN 0-444-41919-5. (Elsevier Scientific: 1981.) $63.50, Dfl.130.

NEELY, W. B. Chemicals in the Environment. Distribution, Transport, Fate, Analysis. Pollution Engineering and Technology, Vol. 13. Pp.245. ISBN 0-8247-6975-9. (Marcel Dekker: 1980.) SwFr.82.

STRANGWAY, D. W. The Continental Crust and its Mineral Deposits. The Geological Association of Canada Special Paper, 20. Proceedings of a Symposium in honour of J. Tuzo Wilson, Toronto, May 1979. Pp.804. ISBN 0-919216-16-1. (The Geological Association of Canada, University of Waterloo, Waterloo, Ontario, Canada: 1980.) Np.

WHITELEY, R. J. (ed.). Geophysical Case Study of the Woodlawn Orebody, New South Wales, Australia: The First Publication of Methods and Techniques Tested Over a Base Metal Orebody of the Type Which Yields the Highest Rate of Return on Mining Investment with Modest Capital Requirements. Pp.610. ISBN 0-08-023996-X. (Pergamon: 1980.) £35.50.

Biological Sciences

ALBANESE, A. A. Nutrition for the Elderly. Current Topics in Nutrition and Disease, Vol.3. Pp.332. ISBN 0-8451-1602-9. (Alan R. Liss: 1980.) $38, Dfl.114.

ALBERTS, B. (ed.). Mechanistic Studies of DNA Replication and Genetic Recombination. ICN – UCLA Symposia on Molecular and Cellular Biology, Vol. 19, 1980. Pp.1003. ISBN 0-12-048850-7. (Academic: 1980.) $48.

BALTES, P. B. and BRIM, O. G. JR. (eds). life-Span Development and Behaviour, Vol.3. Pp.412. ISBN 0-12-431803-7. (Academic: 1980.) $35

BURKE, S. R. The Composition and Function of Body Fluids. 3rd Edn. Pp.208 Pbk. ISBN 0-8016-0903-8. (C. V. Mosby: 1980.) Pbk np.

COHNM, W. E. (ed.). Progress in Nucleic Acid Research and Molecular Biology, Vol.24. Pp.279. ISBN 0-12-540024-1. (Academic: 1980.) $42.50.

EDHOLM, O. G. and WEINER, J. S. (eds). The Principles and Practice of Human Physiology. Pp.672. ISBN 0-12-231650. (Academic: 1981.) £36.80, $88.50.

EISENBERG, J. M. A Collector's Guide to Seashells of the World. Pp.239 ISBN 0-07-019140-9. (McGraw-Hill: 1981.) £12.50.

ERIKSSON, A. W. *et al.* (eds). Population Structure and Genetic Disorders. Sigrid Jusélius Foundation Symposium, Mariehamn, Åland Islands, Finland, August 1978. Pp.690. ISBN 0-12-241450-0. (Academic: 1980.) £55, $126.50.

FREEDMAN, R. B. and HAWKINS, H. C. The Enzymology of Post-translational Modification of Proteins. Vol.1, Molecular Biology. Pp.515. ISBN 0-12-266501-5. (Academic: 1980.) £40, $92.

GORLIN, R. (ed.) Morphogenesis and Malformation of the Ear. Birth Defects: Original Article Series 1980. Vol.16. No.4. Fifth International Workshop held at Gulf State Park, Gulf Shores, Alabama. Pp.367. ISBN 0-8451-1038-1. (Alan R. Liss: 1980.) $44, Dfl.132.

HANCE, R. J. (ed.). Interactions Between Herbicides and the Soil. Pp.350. ISBN 0-12-323840-4. (Academic: 1980.) £20.60, $49.50.

HUNT, R. K. Neural Development, Part II. Neural Development in Model Systems. Current Topics in Developmental Biology, Vol.16. Pp.410. ISBN 0-12-153116-3. (Academic: 1980.) $34.

KANEKO, J. J. (ed.). Clinical Biochemistry of Domestic Animals. 3rd Edn. Pp.832. ISBN 0-12-396350-8. (Academic: 1980.) £60.

MATTA, M. S. and WILBRAHAM, A. C. Atoms, Molecules and Life: An Introduction to General Organic and Biological Chemistry. Pp.721. ISBN 0-8053-9640-3. (Benjamin/Cummings: 1981.) Np.

PHILLIPSON, J. D. and ZENK, M. H. (eds). Indole and Biogenetically Related Alkaloids. Annual Proceedings of the Phytochemical Society of Europe, No.17. Pp.380. ISBN 0-12-554450. (Academic: 1980.) £32.60, $75.

PRICE, J. H. IRVINE, D. E. G. and FARNHAM, W. F. (eds). The Shore Environment. Vol.1, Methods. Vol.2, Ecosystems. The Systematics Association Special Volume, No.17 (a). Proceedings of an International Symposium held at Portsmouth Polytechnic. Vol.1, Pp.321, ISBN 0-12-564701-8; Vol.2, Pp.722, ISBN 0-12-564702-6. (Academic: 1980.) Vol.1 £24.80, $57.50; Vol.2 £48.20, $111.

ROTHMAN, H. *et al.* (eds). Biotechnology. A Review and Annotated Bibliography. Pp.141. ISBN 0-903804-72-7. (Frances Pinter, London: 1981.) £11.

ROTTER, F. L., SAMLOFF, I. M. and RIMOIN, D. L. (eds). Genetics and Heterogeneity of Common Gastrointestinal Disorders. Proceedings of an International Workshop held in Indian Wells, California, March 1980. Pp.582. ISBN 0-12-598760-9 (Academic: 1980.) $35.

SACKNER, M. A. Diagnostic Techniques in Pulmonary Disease. Lung Biology in Health and Disease, Vol.16. Pp.537. ISBN 0-8247-1059-2. (Marcel Dekker: 1980.) SwFr.108.

SCHROEDER, H. E. Differentiation of Human Oral Stratified Epithelia. Pp.308. ISBN 3-8055-1462-X. (Karger, Basel: 1981.) SwFr.159. DM190, $95.25.

SOWERS, J. R. (ed.). Hypothalamic Hormones. Benchmark Papers in Human Physiology, 14. Pp.342. ISBN 0-87933-358-8. (Dowden, Hutchinson & Ross: 1980.) $40. (Distributed by Academic Press.)

STEFFEN, C. and LUDWIG, H. (eds). Clinical Immunology and Allergology. Developments in Immunology, Vol.14. Proceedings of the Symposia at the XIth Congress of the European Academy of Allergology and Clinical Immunology held in Vienna, Austria, October 1980. Pp.420. ISBN 0-444-80312-2. (Elsevier/North-Holland Biomedical: 1981.) $63, Dfl.129.

THODY, A.J. The MSH Peptides. Pp.162. ISBN 0-12-687850-1. (Academic: 1981.) £16, $38.50.

WYNDER, E.L. (Conference Chairman). Plasma Lipids: Optimal Levels for Health. Proceedings of the Conference on Health Effects of Blood Lipids: Optimal Distributions for Populations Held by the American Health Foundation, New York, April 1979. Pp.187. ISBN 0-12-103450-X. (Academic: 1980.) $15.

YOUDIM, M.B.H. and PAYKEL, E.S. (eds). Monoamine Oxidase Inhibitors. The State of the Art. Pp.214. ISBN 0-471-27880-7. (Wiley: 1981.) £13.

ZENZ, C. (ed.). Developments in Occupational Medicine. Pp.477. ISBN 0-8151-9862-0. (Year Book Medical Publishers: 1980.) Np.

ZLOTIN, R.I. and KHODASHOVA, K.S. The Role of Animals in Biological Cycling of Forest — Steppe Ecosystems. Pp.221. ISBN 0-87933-377-4. (Dowden, Hutchinson & Ross: 1980.) $19.50. (Distributed by Academic Press.)

ANNOUNCEMENTS

Meetings

15-16 June, **Safety and Loss Prevention**, York (Mrs G.M. Nelson, IChemE, George E. Davis Building, 165-171 Railway Terrace, Rugby, UK).

16-17 June, **The Immune System**, Basel (Basel Institute for Immunology, Hochhaus, Bau 52, Switzerland).

16-18 June, **Space Shuttle**, Wembley (The Institute of Measurement and Control, 20 Peel St, London W8, UK).

18-19 June, **Chromatography in Biochemistry, Medicine and Environmental Research** *and* **Mass Spectrometry in Biochemistry, Medicine and Environmental Research**, Venice (Dr A. Frigerio, Via Eritrea 62, 20157 Milano, Italy).

23-24 June, **Food and Health — a Perspective for the 80's**, London (The British Nutrition Foundation, 15 Belgrave Square, London SW1, UK).

24-25 June, **Biotechnology: A Briefing for Scottish Industry**, Strathclyde, (Dr J. Aitchison, University of Strathclyde, Royal College, 204 George St, Glasgow, UK).

22-25 June, **Erosion and Sediment Transport Measurement**, Florence (Dr P. Tacconi, Istituto de Ingegneria Civile, Via S. Marta 3, 50139 Firenze, Italy).

25-26 June, **Genetic Engineering in Food and Agriculture**, Virginia (R. Nash, The Energy Bureau Inc, 41 East 42nd St, New York, New York 10017, USA).

30 June-10 July, **Mechanical and Thermal Behaviour of Metallic Materials**, Varenna (Societa Italiana di Fisica, Via L. Degli Andalo, 2, 40124 Bogna, Italy).

2 July, **Health Hazards of VDUs?**, Loughborough (B. Pierce, HUSAT Research Group, Dept of Human Sciences, Loughborough University of Technology, Leicestershire, UK).

3 July, **Developments in Current Polymer Research**, Aberdeen (Prof. P. Meares, Dept of Chemistry, University of Aberdeen, Meston Walk, Old Aberdeen, UK).

6-10 July, **An Introductory Course in Tropical Hygiene**, London (Ross Institute of Tropical Hygiene, London School of Hygiene and Tropical Medicine, Keppel St, London WC1, UK).

6-10 July, **Mixing in the Process Industries**, Bradford (Dr N. Harnby, Dept of Chemical Engineering, University of Bradford, West Yorkshire, UK).

6-10 July, **Process Safety—Theory and Practice**, Durham (Mrs M. Findlay, Dept of Chemical Engineering, Teesside Polytechnic, Borough Rd, Middlesbrough, Cleveland, UK).

7-9 July, **Light Microscopy '81**, London (Administrator, Royal Microscopical Society, 37/38 St Clements, Oxford, UK).

29 June-10 July, **Summer School in Health Physics**, London (Dr H.D. Evans, Imperial College, London SW7, UK).

9-10 July, **Cost Savings in Distillation**, Leeds (Prof. G.G. Haselden, Dept of Chemical Engineering, University of Leeds, Leeds, UK).

13-16 July, **The Technology of Fibre-reinforced Plastics Composite Materials**, Guildford (Mr M.G. Bader, Dept of Metallurgy and Materials Technology, University of Surrey, Guildford, UK).

13-18 July, **Applied Rock Mechanics**, Davis (Garrett Jones, University Extension, University of California, Davis, California 95616, USA).

14-16 July, **Inter/Micro — 81**, Cambridge (McCrone Research Institute Ltd, 2 McCrone Mews, Belsize Lane, London NW3, UK).

14-24 July, **Positrons in Solids**, Varenna (Societa Italiana di Fisica, Via L. Degli Andalo, 40124 Bologna, Italy).

16-17 July, **Breeding Legumes for Nutritional Value**, Southampton (Nutrition Society, Dr M.I. Gurr, Shinfield, Reading, UK).

19-24 July, **Chemistry of Platinum Group Metals**, Bristol (Dr J.F. Gibson, The Royal Society of Chemistry, Burlington House, London W1, UK).

21-23 July, **Synthesis in Organic Chemistry**, Oxford (Dr J.F. Gibson, The Royal Society of Chemistry, Burlington House, London W1, UK).

22-24 July, **HPLC Beginners Course**, Cheshire (HPLC Technology Ltd, 4 Charlotte St, West Macclesfield, Cheshire, UK).

28 July-2 August, **Data Acquisition in High Energy Physics**, Varenna (Società Italiana di Fisica, Via L. Delgi Andalò, 2, 40124 Bologna, Italy).

17-21 August, **Earth Tides**, New York (Prof. J.T. Kuo, 828 S.W. Mudd, Columbia University, New York, New York 10027, USA).

17-21 August, **Soil Morphology**, London (Dr P. Bullock, Rothamsted Experimental Research Station, Harpenden, Herts, UK).

23-28 August, **Solar World Forum**, Brighton (UK-ISES, 19 Albemarle St, London W1, UK).

28 August-1 September, **9th International Congress of ISDB**, Basel (A.A. Moscona, University of Chicago, Dept of Biology, Chicago, Illinois 60637, USA).

28 August-9 September, **Arc Volcanism**, Tokyo (Prof. D. Shimozuro, Earthquake Research Institute, University of Tokyo, Bunkyo-Ku, Tokyo 113, Japan).

30 August-4 September, **Regulatory Peptides: Functional and Pharmacological Aspects**, Brescia (M. Trabucchi, Dept of Pharmacology and Therapeutics, Via Valsabbina, 19-25100 Brescia, Italy).

30 August-4 September, **Optics in Biomedical Sciences**, Graz (Medizinische Ausstellungs und Werbegesellschaft, M. Rodler & Co. A-1014 Wien I, Freyung 6, Postfach 155, Austria).

31 August-4 September, **BAAS**, York (BA, Fortress House, 23 Savile Row, London W1, UK).

1-5 September, **Central Nervous System Regeneration**, Catania (Dr A.M. Giuffrida Stella, Istituto di Chimica Biologica, Universita di Catania, Viale Andrea Doria, 6, 95125 Catania, Italy).

2-4 September, **Environment and Safety**, Wembley (IE & S Exhibition and Conferences, Newgate, Sandpit Lane, St Albans, Herts, UK).

31 August — 11 September, **Eighth UICC Training Course in Cancer Research**, Heidelberg (Prof Dr G. Sauer, Grman Cancer Research Center, Im Neuenheimer Feld, 6900 Heidelberg, FRG).

1-4 September, **2nd Australian National Laser Conference**, Canberra (The Secretary, 2nd Australian National Laser Conference, Dept of Engineering Physics, Research School of Physical Sciences, PO Box 4, Canberra, ACT, 2600, Australia).

2-4 September, **Modern Aspects of Phase Equilibria**, Teesside (Mrs M. Findlay, Dept of Chemial Engineering, Teeside Polytechnic, Borough Rd, Middlesbrough, Cleveland, UK).

5-10 September, **Precision and Speed in Close Range Photogrammetry**, York (K.B. Atkinson, Dept of Photogrammetry and Surveying, University College London, Gower St, London WC1, UK).

7-11 September, **Particle Size Measurement**, Bradford (Secretary, School of Powder Technology, University of Bradford, Bradford, West Yorkshire, UK).

7-11 September, **Optics in Biomedical Sciences**, Graz (Medizinische Ausstellungs-und Werbegesellschaft, M. Rodler & Co. A-1014 Wien 1, Freyung 6 Postfach 155, Austria).

7-11 September, **Modelling and Simulation of Dynamical Chemical Engineering Processes**, Bradford (Dr J. Ingham, School of Chemical Engineering, University of Bradford, Bradford, West Yorkshire, UK).

8-11 September, **Down to Earth Management**, Winnipeg (7th Canadian Symposium on Remote Sensing, PO Box 1106, Winnipeg, Manitoba, Canada R3C 2X4).

10-11 September, **Molecular Beams in Chemical Physics**, Birmingham (Dr A.R. Burgess, Dept of Chemical and Biochemical Engineering, University College London, Torrington Place, London WC1, UK).

13-18 September, **15th International Congress of the International Society for Chronobiology**, Minneapolis (ISC XV Conference, College of Pharmacy, 5-110 HS Unit f, 308 Harvard St SE, Mpls, Minnesota 55455, USA).

17-20 September, **British Veterinary Association's Annual Congress**, Exeter (BVA, 7 Mansfield St, London W1, UK).

nature

11 June 1981

What future for postdoctoral fellows?

Are postdoctoral fellows the lowest form of academic life? This belief, common enough, is well supported by the evidence now gathered by the Grodzins committee (see page 443). Postdoctoral fellows are badly paid by comparison both with PhD graduates who have chosen other careers and, more galling, with those already lucky enough to have found permanent academic posts. The conditions of employment — sometimes the lack of medical insurance or of retirement benefits — are often demeaning, yet postdoctoral fellows shoulder a disproportionate share of responsibility for academic research programmes which permit the building of an independent reputation only when circumstances are right. So why do 12,500 scientists in the United States put up with these indignities? Classically, the prize has been a tenured position at a good university. The chance of that is no longer high, at least in the more crowded fields (biology in particular). So postdoctoral fellows have a growing sense, often just, of being exploited.

So is this scandal? The Grodzins committee rightly shrinks from giving the simple answer yes to this oversimple question. To be sure it would be outrageous, if true, that self-governing bodies of academics — American universities — should deprive their temporary colleagues in research of medical insurance and the like. But the American system of postdoctoral fellowships is different both in conception and in its present function from some of the systems in place elsewhere. The intention, adapted from the old German system of the *Privatdozent,* is that postdoctoral fellows should be apprentice academics — people of promise who are blessed with a few years in which to establish independent reputations as scholars. At the turn of the century in Germany, such people were paid less well than established academics, chiefly because the universities were chronically short of funds. In the United States, the system has never been strictly comparable with the German, if only because American universities have mostly relied on outside sources for funds with which to support the bulk of postdoctoral fellows. In the past few decades, the system has been further unsettled by rapid expansion, especially in biology, and by the relative decline of academic posts. The avoidance of all disappointment is impractical. What else should be done?

The most urgent need is that universities should take more direct responsibility for their postdoctoral fellows. If the implicit promise of such appointments — a tenured teaching post — has to be broken, the university cannot decently shrug its shoulders and let the disappointed fellow walk away. Although people now embark on fellowships like these with their eyes wide open to the risks, after three years or so and the accretion of personal as well as professional responsibilities, a disappointed fellow needs more than an expression of sympathy. (Pay during the tenure of a fellowship matters less.) The most constructive if the most nebulous of the recommendations of the Grodzins committee is that universities should establish continuing machinery within which recurring problems can be discussed. The underlying need is that universities should acknowledge both the dependence of their reputations in research on the floating population of postdoctoral people and the need that, if a stint should end in disappointment, there should be some tangible means by which a postdoctoral fellow can thereafter make his way in the world.

The problem is not merely a problem for universities in the United States. Over the past several years, in most of the university systems in which research is (in practice, as well as theory) an important part of everybody's work, it has turned out to be expedient that PhD graduates should be the backbone of the research effort in many academic departments. Broadly speaking, they are able (intelligent and skilled) and willing. Not all people with these qualifications are formally included in the category of postdoctoral fellow used by the Grodzins committee, which technically excludes those whose salaries are provided by short-term research contracts but also, for example, people elected with considerable solemnity to junior fellowships at Harvard University and their equivalents elsewhere.

In countries such as France, the universities in partnership with organizations such as the Centre National de la Recherche Scientifique have sought to solve the difficulties which arise by means of essentially permanent appointments to research posts in parallel with regular academic posts. The snag is that too few young people enter research, and that dead wood accumulates. Elsewhere, in Britain for example, the embarrassment caused by the frequent need to tell people living off soft money that there will not be an extension has led to the convention among grant-making agencies that after one or at the most two stints, postdoctoral workers should be told that, however talented, they are ineligible for further support; this is a device for making the problem go away, not for solving it. The American practice, more tolerant of repeated extensions of postdoctoral fellowships, has the effect of postponing the same problem. The risk is that the brightest people will be among the first to leave.

The Grodzins committee can in the circumstances be forgiven for having very little to offer as a remedy for the circumstances that have arisen in the United States. Nowhere have the problems of postdoctoral employment been dealt with easily. The committee's proposal that there should be a system of federally-financed postdoctoral fellowships in the original mould of academics-in-waiting would help, but not much. So long as the universities need postdoctoral fellows to assist with their research programmes, and for as long as grant-making agencies consider that posts of this kind provide valuable post-postdoctoral training — and can afford to pay the stipends — the creation of new fellowships will continue. The dramatic expansion in the biological sciences in the past few years will and should continue, for there is plainly no end in sight to the biological projects that can profitably be undertaken. Almost by accident, however, the new Administration's budget may change the shape of the expansion. By outlawing the contributions the National Institutes of Health have been making to the universities receiving postdoctoral fellows, the Administration has probably ensured that more postdoctoral fellowships will be paid out of project grants, not from a dedicated budget. That will at least help to reduce unreasonable expectations of continued academic work.

The other steps that need to be taken are more practical, and humdrum. Even when postdoctoral fellows have embarked on their careers knowing what the risks may be, their eventual departure from the university may be a hardship, even an injustice. The job market in the field concerned may have changed, or the point at which a contract ends may be one at which it is exceedingly improbable that a suitable job elsewhere might be found. So is there not a case for asking that universities (rather than the sponsoring agencies) should be able to soften the problems of this inevitably painful transition period? Several stratagems are possible. The ideal would be that postdoctoral fellows at the end of their contracts and unable then to find a suitable job should be paid a bounty, a kind of retrospective recompense for the deprivations of earlier years. At the Massachusetts Institute of Technology, a scheme for providing

0028-0836/81/240441-02$01.00

short but limited spells of continued employment (at the expense of the institution) is being considered. Neither of these schemes is ideal. All schemes for easing the transition from academic life to the outside world will leave much to be desired (for not all disappointment is avoidable) — and will cost money. But the Grodzins committee shows clearly enough that something needs to be done. It is unseemly that the academic community, while rightly insisting on the highest standards of scholarship and achievement from candidates for permanent faculty positions, can as casually as at present shrug off its responsibilities for those who, for one reason or another, fail to be chosen.

Change wanted

The University of London, in the throes of reorganization or at least of the contemplation thereof for the past year, will not be much helped by the second interim report from the Swinnerton-Dyer Committee, which has been charged with suggesting how student numbers might be reduced and the cost of operations still more rapidly reduced. For what the committee says (see page 447), while sensible and moderate enough, provides neither a prescription with which the several components of the university will be unable to argue nor a means by which they will sense an incentive to reorganize themselves. The university has a long, disputatious summer ahead of it.

What is happening in London is nevertheless a foretaste of what is likely to be happening in other British universities a few weeks from now, when the University Grants Committee translates into putative grants for individual universities the dire threats it has been making of hard times ahead. The university differs from the other universities in Britain, except the University of Wales, in being a federal institution. For much of the past fifteen years, the various London colleges have become more and not less autonomous as the federal centre has shed responsibility for extraneous tasks such as the supervision of academic standards in associated institutions (many of which have become universities in their own right) and the provision of public examinations in much of the British Commonwealth. The separate colleges of the university with responsibility for teaching undergraduates are now almost autonomous in awarding degrees. All of the colleges and most of the specialized institutes of the university employ their own members of staff. With the exception of Imperial College, which has a direct relationship with the University Grants Committee, the several components of the university are financed by means of grants from what is called the "court" of the university, which divides whatever it can spare according to rules and formulae which are not made known, even to the recipients of its cheques.

The prospect, as described by the committee, is daunting. In the course of the three years ahead, income will be reduced by 15 per cent and the committee has taken the sensible view that economizing on equipment and other materials has already gone far enough. So the threatened cut implies smaller payrolls, reducing more quickly than natural wastage will permit. With that starting point and the implicit assumption that academic staffs and their ancillaries will be reduced in equal proportions, the committee arrives naturally at its conclusion that at least one college must be disbanded, and that there should be a system of peer review for deciding which academic departments elsewhere deserve the university's support. This is what people within the university will be arguing about. Why Chelsea College for oblivion? Why not somewhere else? And so one.

The university will be a less wretched place in the months ahead if it can concentrate instead on the more constructive parts of the committee's report. On the proper grounds that it would seem desperately unfair that all teachers of say, psychoastrometry in one of the London colleges should lose their jobs because the process of peer review had decided that the teaching of this subject should be concentrated elsewhere within the university, the committee argues that there should be an attempt to make the university rather than its separate parts the ultimate employer of all the university's academics. Then, the argument goes, it would be possible to ensure that even if it seems sensible to concentrate the teaching of psychoastrometry in some larger department of astrometry on some other site, psychoastrometry will still be taught by the best psychoastrometrists. The argument is unshakable while, in their present panic, the university's academics would probably agree that their contracts of employment should be transferred to an employer of the last resort, the university.

The university must also grasp the more painful nettle of deciding which kinds of employees it hopes to keep on its books. The assumption that academic and non-academic staff must be reduced in step is genteel but not essential. The peer review of academic performance that the Swinnerton-Dyer Committee asks for will take time, and the outcome cannot be conclusive. A managerial review of the necessity of the armies of ancillary staffs employed by the university could be quicker and might prove a more effective economy. It would naturally be wrong to think that an academic institution such as the largest university in the United Kingdom could be run entirely by its academics, but is there not a case for thinking that fewer cleaners, porters and even technicians might be prudent in parlous times? The academics in the University of London, as elsewhere, will shrink from this question, some of them because they cannot imagine how their own energies might be bent to deciding what denomination of stamp should be stuck to what kind of letter, most of them because they would consider the enquiry inequitable as between academics and others. But is not the university an institution whose chief purpose is academic?

Some silver lining

In the United States as elsewhere except France — for the time being — professional people are adjusting to smaller public subventions for research. The cause is everywhere the same: inflation. By now, it is well-established that the threat posed by inflation to civilizing institutions (universities, research laboratories, even entrepreneurial businesses) is unavoidable even if, for a time, insidiously unrecognized. Occasionally, however, some casualties of the attempts by governments to offset inflation by reducing spending are not losses at all. So much is illustrated by the new Reagan budget, widely and justly bemoaned in the United States. Although the new Administration appears to have acted over the whole range of public policy without stopping to calculate the consequences (in which it resembles the Thatcher government's policy towards British universities) some of the uncalculated consequences will be welcomed. The following are some examples.

President Carter's budgets for the National Science Foundation has been left unchanged by President Reagan except in one respect — the foundation's educational budget was cut by a half, that for the support of the social sciences virtually eliminated and that for international activities reduced. Tactically, the Administration has been clever. Only a small part of the foundation's constituency will have been hurt so badly that it squeals aloud. The threatened areas are also those in which the National Science Foundation has been living dangerously. Nobody can now claim that the curriculum development the foundation has been sponsoring is pointed and effective; rather, too little resources are being spread too thinly. And although it is unforgivable (and will not be forgiven in Europe) that the Administration should have cut the foundation's international budget without prior consultation, there is at least some merit in asking (as many in Washington are doing) whether continued support for the International Institute for Applied Systems Analysis in Vienna, to the tune of nearly $3 million a year, is the best use for such a sum of money. None of this implies that the Reagan cuts are to be welcomed for the hardship they will bring and the damage they will do. The moral, if there is one, is that if times ever improve, people should discriminate more carefully between good projects and the not so good.

0028-0836/81/240442-01$01.00

Anxiety about postdoctoral prospects

Academy report sees trouble in store

Washington

A postdoctoral position may be a necessary but is no longer a sufficient stepping stone to a tenured academic position in the United States, according to a report now published by the National Research Council, the research arm of the National Academy of Sciences. Ironically, the council was one of the first agencies to operate a system of postdoctoral fellowships soon after its own creation in 1916.

The report, commissioned by the National Science Foundation and prepared by a committee under Dr Lee Grodzins of the Massachusetts Institute of Technology, starts from the view that postdoctoral fellowships make an invaluable contribution to research, and remain a valuable preparation for academic life. It is dismayed by the mismatch between the growth in the 1970s of the numbers of PhD graduates employed on postdoctoral grants and the decline of openings in tenured faculty positions.

In 1979, there were an estimated 12,500 postdoctoral fellows working in scientific fields, some 5 per cent of the total doctoral labour force in United States universities and an increase of 50 per cent compared with 1972. Roughly a fifth of postdoctoral positions in 1979 were in government laboratories and other such institutions.

The biosciences forced the pace of expansion in the 1970s, accounting for 7,300 postdoctoral fellows in 1979. The committee cites federal policies as the driving force and the increasing proportions of new PhD graduates seeking postdoctoral positions (62.7 per cent in 1979) and the increased duration of postdoctoral stints as ancillary causes. By contrast, the numbers of postdoctoral fellows in physics and chemistry (1,100 and 1,700 respectively in 1979) are declining both in absolute numbers and as proportions of all postdoctoral fellows or of those obtaining PhDs in physics and chemistry. In engineering and mathematics, postdoctoral fellows in 1979 numbered only 400 and 200 respectively, a mere 1 per cent of the total numbers in academic posts (not all tenured).

The career benefits of a spell as a postdoctoral fellow are mixed. Data for the employment of the 1972 cohort of PhDs suggest that graduates make an early choice between a postdoctoral stint at one of the 59 "major research universities" (defined as the institutions with two-thirds of all research budgets in higher education) and a tenured post elsewhere. Thus only 5 per cent of those graduating in 1972 in the biological sciences and electing for a postdoctoral stint had found tenured positions at research universities, but 28 per cent of those deciding otherwise had by then found tenured positions elsewhere (while 9 per cent had found posts at the major universities).

Hiring practices vary considerably from one field to another. In the major research universities, 88 per cent of the most recently appointed assistant professors in chemistry had done a postdoctoral stint, compared with 82 per cent in physics, 73 per cent in the biological sciences and only 29 per cent in engineering.

There is some evidence that a postdoctoral stint can mean a lasting loss of salary. Even in tenured faculty posts, the median salaries of those with postdoctoral experience (from the 1972 cohort) are between $1,000 and $2,000 less than those of direct entrants. The differences are even larger for those who have found jobs in government service and industry. Most people seem to have had to extend their postdoctoral stints by at least a year while seeking suitable jobs, and there is now some evidence that new PhDs are less inclined to look for postdoctoral appointments — the percentages of PhD graduates seeking postdoctoral appointments have been declining since the early 1970s in chemistry and physics, and appear to have reached a plateau (above 60 per cent) in the biological sciences.

The report resembles others in its evidence that women and members of minority groups are at a disadvantage both in finding postdoctoral appointments and in profiting from them. In 1979, overseas scientists accounted for 38 per cent of all postdoctoral appointments and for 70 per cent of postdoctoral appointments in engineering. The report says nothing about the numbers of US citizens holding postdoctoral positions overseas.

Low pay, indignities and uncertainty

Washington

Among postdoctoral fellows in the United States, those working in the biological sciences are the most numerous, the most eager and the most at the mercy of the job market. Of PhDs in the biological sciences, 58 per cent take on a postdoctoral stint (compared with 50 per cent in physics). Half of biology postdoctoral fellows hold their appointments for two years or more and 43 per cent of them have to prolong their stints while looking for jobs. (Both chemistry and physics are under more pressure, with 38 per cent and 36 per cent respectively.)

Postdoctoral stipends are lower in biology than in other fields except psychology and the social sciences, at an average of a mere 57 per cent of tenured academics of the same seniority. (Mathematician postdocs can expect to earn 96 per cent of an academic salary.) Biologists on fellowships also make up the largest proportion in their field of all research personnel (16 per cent) and of equivalent full-time research staff (27 per cent) in major research universities.

The Grodzins report, *Postdoctoral Appointments and Disappointments* (National Academy of Sciences, $17.75), justifies the second part of its title with anecdotal as well as statistical information gathered by a questionnaire to graduating PhDs. Postdoctoral stipends seem to have been a constant source of complaint. One said that the postdoctoral position he turned down seemed to be "little more than a glorified graduate student", another that postdoctoral fellows are not "treated as professionals", earning as they do less than "laboratory technicians, some secretaries, janitors, etc".

There appears to be a strong feeling among biology postdoctoral fellows that the "pay-back" provisions of fellowships awarded by the National Institutes of Health are iniquitous. NIH fellows are required to refund part of their stipends if they do not match the duration of their fellowships with an equal time in some other health service field. One fellow compared this condition with "indentured service", another (a woman) pointed out that the requirement would conflict with her family life.

Many postdoctoral fellows acknowledged the opportunities provided for getting on with personal research, sometimes as a prelude to a complaint of having been "exploited" as "second class citizens". The lack of medical insurance and retirement provision appears to be resented as demeaning. The report quotes an eloquent appeal from a mathematician for more but more rewarding postdoctoral fellowships in his field, and an engineering professor notes that it is now almost impossible to persuade American citizens to take fellowships because of the high salaries in private industry.

Many respondent PhDs told the committee that they regarded postdoctoral fellowships as a "holding pattern for the academic life" — but the committee says there is no evidence that universities are losing the brightest people because of the uncertainties and indignities involved. Employers outside the university sector differ in their estimations of the value of postdoctoral experience — government laboratories welcome postdocs, private industry sometimes finds them inflexible.

0028-0836/81/240443-02$01.00

On the grounds that postdoctoral positions are increasingly being used to support academics in waiting, the Grodzins committee urges that the federal government should establish a system of postdoctoral awards (perhaps 250 new awards a year) at stipends half as much again as awards now available, that there should be roughly 50 similar awards for overseas scientists and that universities should give more thought to the careers of their postdoctoral fellows, setting up oversight committees for the purpose. Fearing employment difficulties for postdoctoral people, the committee also asks that the National Science Foundation extend the scope of its present demographic work to cover postdoctoral fellows.

US minerals

Resource warcries

Washington

Is the Soviet Union conducting an undeclared "resource war" against the West in an effort to jeopardize supplies of minerals needed for high technology products such as computers and military aircraft? Yes, according to conservative foreign policy groups such as the Washington-based Center on Economics and National Security, claiming that recent Soviet moves to increase purchases of so-called critical minerals on the international market is part of a broader strategy to weaken the West's military defences.

Not so, according to a group of geologists, mining engineers and assorted Soviet watchers who met last month at the Johns Hopkins University's School of Advanced International Studies. Although agreeing that the Soviet Union had, indeed, stepped up its purchases of critical minerals and reduced its exports, the meeting found little evidence that this was part of a deliberate military strategy, but rather reflected the Soviet Union's difficulties in meeting its own mineral needs.

The debate is not merely academic — for example the concept of a "resource war" is already being used by legislators such as Representative Jim Santini of Nevadah to justify attempts to open up federal lands in the United States for exploration by mining companies, a move firmly opposed by environmentalists. Similar arguments are being used to the same ends by the new Interior Secretary, Mr James Watt.

There is little debate on the principal facts of mineral supply. The United States is almost totally dependent on foreign countries for many of the minerals used in high technology products, such as the manganese necessary for specialty steels (of which South Africa is the main supplier), and the cobalt needed in jet engine alloys (93 per cent of which is imported from Zambia and Zaire).

In theory, this makes the United States highly vulnerable to cartel-type actions such as those imposed by the oil-producing countries in the early 1970s. In practice, policy-makers are having to decide how likely such actions might be, and what could be done to minimize the chance of dislocation, for example by developing substitute materials.

The new Administration has started on a much more hawkish note than its predecessor. Secretary of State Alexander Haig, as president of United Technologies, had previously talked of "resource battlegrounds" in the Third World. And soon after his election last November, Mr Reagan set up a task force on strategic minerals chaired by Mr Daniel McMichael, president of the World Affairs Council of Pittsburgh, which had done much to legitimize the "resource war" concept.

Last June, for example, the World Affairs Council organized a conference in Pittsburgh on this theme, discussing in particular a report from Dr Daniel Fine, a resource analyst affiliated to the Massachusetts Institute of Technology's Mining and Mineral Resources Institute. Dr Fine argued that recent changes in Soviet minerals policy had conscious strategic implications; for example, that a decision to halt the export of titanium sponges to the United States "was based on an apparent decision to deter the US aerospace defence build-up forecast for the early 1980s".

The task force, which spurred the new Administration's general approach to minerals policy, also drew heavily on members of the Center on Economics and National Security, which had issued similar warnings of Soviet intentions in a report published last June. This committee, which receives support from the same conservative foundations as the Pittsburgh group, helped to draft a statement issued by the American Geological Institute last October warning the three main presidential candidates of Soviet efforts "to deprive this country of strategic minerals indispensable to its national defence".

Those who might be described as moderates on the strategic minerals issue accept its political importance. But they argue that major difficulties can probably be avoided by both renewed stockpiling and greater efforts to develop substitutes (which are beginning to go under the name "synterials"). Some suggest that, if there is going to be a "resource war", then it is more likely to be the result of competition with other customers, such as Japan.

The hard-liners push for a more aggressive stance. With the "national security" arguments playing a more dominant role, much of this push is now likely to come from within the Administration. One casualty of this thinking may be the United Nations Law of the Sea — the Reagan administration is concerned that the present UN proposals would not guarantee the United States access to sea-bed supplies of manganese and cobalt. **David Dickson**

National Cancer Institute

Congress complains

Washington

Is Congress going sour on the conquest of cancer? This is the question raised in many minds by the tough questioning last week of the director of the National Cancer Institute (NCI), Dr Vincent T. DeVita, by the Senate Labor and Human Resources Committee. At one point, DeVita protested "I'm not a fool. I want this committee to be patting me on the head, not on other parts of my anatomy".

For the past decade, the committee has been one of the staunchest supporters of cancer research, helping to bid up the budget of NCI to more than $1,000 million. Last week, however, the committee's hearing was delayed by an internal wrangle. Senator Edward Kennedy, senior minority member of the committee, complained that a dossier of complaints against NCI had been leaked in advance to the *New York Times* and the *Washington Post* without having been shown in advance to him and his Democratic colleagues. There was further dispute over the proposal by the committee chairman, Senator Orrin G. Hatch, that he should introduce each section of the questioning with a series of previously prepared questions. Even with the new low-power television lights, some senators perspired.

The chief bone of contention last week was NCI's award of a grant of $910,000 to Dr Marc J. Strauss early in 1980 although Dr Strauss had resigned from his previous position at Boston University after allegations of falsified results two years earlier. A report on the allegations by the National Institutes of Health is expected later this summer, although the committee also had access to a report of the Food and Drug Administration which purports to document the mismanagement and even falsification of patient records as part of Boston University's work for the North-Eastern Cooperative Oncology Group.

The committee and Dr DeVita agreed last week on a version of the events at Boston University. Doubts about the adequacy of patient records arose within Boston University in June 1978, and Dr Strauss lost his job in August of that year. Since then data from the Boston studies have not been pooled with other data.

A year later, Dr Strauss was said to have applied for a further grant of $910,000. Dr DeVita became director of NCI in January 1980, and agreed last week that he had deliberately decided not to inform NCI's advisory board of the allegations against Dr Strauss because the new application did not entail the treatment of human patients, because the case against him had not been proved and because Dr Strauss must be considered innocent until proved guilty.

The allegations said by Senator Hatch to have been justified by the Food and Drug Administration report include the use of doses of ^{3}H-thymidine in excess of those required by the agreed protocols in the treatment of cancer patients; failure to obtain the informed consent of 112 patients to an experimental treatment; the treatment of 27 patients with ^{3}H-thymidine for breast cancer which was not one of the conditions specified in the protocol; and failure to keep complete X-ray records. Dr DeVita agreed that "all the allegations are curious and, if true, reprehensible".

The committee asked three broad questions. Why had the NIH investigation taken three years to bring nearly to completion? Why had Dr Strauss not been denied a further grant simply on the grounds that serious allegations had been made against him? And what would happen if the allegations were substantiated?

Dr DeVita admitted that "we were too slow off the mark" and also that "with the benefit of hindsight" it would have been wise to have told the advisory committee (a few weeks after becoming director of NCI) that Dr Strauss's Boston record was being investigated, even though that would probably have meant that the $910,000 grant would not have been made. But he insisted that his discomfort before the committee was matched by his discomfort at "the idea of judging someone guilty until the investigation is complete".

Meanwhile, Dr DeVita said, procedures at NCI have been tightened up, with contracts reviewed by a central financially oriented staff. All contracts will in future be subject to peer review and he personally will take responsibility for seeing that funds are properly spent. If the allegations against Dr Strauss are proved, it seems there will still be time to recover $600,000 of the 1980 grant.

US astronomy

Prospects for people

Washington

Anxiety about the future of astronomy in the United States is one of the recurrent themes in the final draft of the National Research Council report of the Astronomy Survey Committee. The committee, whose chairman is Dr George B. Field from the Harvard/Smithsonian Center for Astrophysics, has now submitted its fifth draft to the review procedures of the National Academy of Sciences.

The final draft is especially concerned about the plight of those astronomers who have been employed on fellowships or whose salaries have been charged to project grants for periods as long as six or eight years, and whose careers are now threatened by declining budgets. Astronomy differs from other fields of science in that there are no natural industrial outlets for academic astronomers, while universities are often unwilling as well as unable to employ such relatively senior people.

The draft report recommends a system of assistant professorships financed jointly by universities and the federal government. The acceptability of this proposal, for which there is no precedent in the United States, has yet to be determined.

The Field Committee also plans a series of recommendations of priorities for the future development of astronomy. Predictably it puts its weight behind the space telescope, but urges that support should also be given to second-generation instruments for the device. The launching of a γ-ray observatory (by NASA) and the building of the proposed 25-metre millimetre-wave telescope are in the list of projects to be pursued, and there is a cautious welcome for the Princeton proposal to construct a gallium detector for solar neutrinos.

Further ahead, the Field Committee apparently wishes to give priority to an X-ray observatory even more advanced than Einstein and to a "new technology" telescope for optical and infrared observations. The proposal to deploy a 10-metre dish in an orbit about the Earth is put next on the list. The draft report has much to say about the need for technical and supporting services for the development of these projects.

Publication of the report is unlikely before October.

Vive la différence!

M. Pascal Clément would like French scientists to speak more French, *s'il vous plait*. In his report on the state of the French language, commissioned by the then President Giscard d'Estaing and presented recently to President Mitterand, M. Clément — a parliamentarian and member of Giscard's Union pour le Democratie Française — devotes a whole chapter to the invasion of English into science in France. French scientists are lazy about translating English technical terms, says Clément, preferring a kind of "Franglais"; and they even write their papers in English.

Something must be done, the report demands. Clément suggests that conferences conducted in English without translation facilities should get no support from state funds. This — if it ever took effect — would worry the Centre National de la Recherche Scientifique which supports much of the basic science in France; for two-thirds of papers read at its conferences are read in English, while only one in five conferences have translators present.

Scientists have reacted to Clément's strictures with amused smiles. He is naive, they say. Much as they would like to speak and write in their native tongue, French scientists constitute only a tiny fraction of the world's scientists (7 per cent, estimates one). And any scientist worth his or her salt in France will be collaborating with groups abroad, with half of them probably American. And how many Americans can read or speak French?

Moreover, the inflexible attitude of people like Clément towards their language probably does more to help English penetrate French than hinder it, they say. For this attitude sees mistakes of foreigners attempting to speak French as *fautes* — faults — in the speaker, thus embarrassing some who would otherwise try to get by in the language.

As for the consequences of the report on French science, the thing is *une lettre morte* already. **Robert Walgate**

Uranium enrichment

US price hike?

Washington

The question whether the United States should charge customers for enriched uranium more than the bare cost of enrichment surfaced again last week in the House of Representatives, this time because Congress is looking for ways of raising extra revenue so as to relieve the budgetary dilemmas forced on it by the Administration. But the chances of raising extra money quickly are small. And the Administration is clearly divided within itself about the merits of the proposal.

The charge for enrichment services at government enrichment plants has been a controversial issue since 1974, when it was planned that responsibility for non-military enrichment should be transferred to the private sector, and when everybody agreed that government prices should not undercut those of private manufacturers.

Last week, Congressman Richard L. Ottinger, chairman of the House Subcommittee on Energy Conservation and Power, opened a brief and hastily arranged enquiry on the subject with the blunt confession that the House Budget Committee needed to know this week whether the price of enriched uranium should be increased. One of his difficulties was that the Department of Energy, which argued for an increase of price in 1978, has now changed its mind. Dr William R. Voigt, who put the department's case for an increase three years ago, last week argued that the rules for enrichment pricing should not be changed simply so that Congress could spend more on other energy projects.

The General Accounting Office has been more consistent. In April, the director of the Office, Dr Dexter Peach, argued for prices that would take account of what private companies would have to set aside for state and federal taxes and for making a return on capital. His April calculation, repeated last week, was that the US enrichment business, now worth $1,300 million a year, would yield an extra $1,300 million

0028-0836/81/240445-02$01.00

between now and October 1986, one-third of that from overseas.

The nuclear industry (represented last week by the Edison Electric Institute) objects to the proposed increases of price partly on legal grounds — the amended Atomic Energy Act would have to be amended again — and partly on the grounds that a change in the price structure for enrichment services should be part of a more general review.

The costs of enrichment are already due to increase by 18 per cent from October. The proposed "fair pricing" increase would amount to a further 15 per cent, implying an increase of the cost of electricity generated by nuclear stations of between 1 and 2 per cent. Such an increase would not be decisive, which is why the decision on fair pricing may eventually turn on the competitiveness of US enrichment. Dr Peach quoted last week the latest prices charged by Eurodif as $127 for a unit of separative work. For the first time, US enrichment is not the world's cheapest.

UK nuclear power

Missed again

The ailing nuclear construction industry in Britain, which suffered a ten-year gap in orders in the 1970s and over-committed itself to the advanced gas-cooled reactor, may now even be denied a quick entry to American pressurized water technology. For orders for key parts of the proposed Sizewell B nuclear power station — to be the first commercial pressurized water reactor (PWR) in the United Kingdom if it survives a public inquiry — may have to be placed abroad.

The reasons are cost and time. British industry is not tooled up to produce PWR components, in contrast with other companies — notably those in France, the United States and Germany — so costs will be higher and delays longer if orders are placed in the United Kingdom. Of the principal possible UK contractors, Babcock and Wilcox has indicated that it would need a government subsidy to compete with foreign suppliers; and Northern Engineering Industries has asked that foreign competition be ruled out.

Neither now seems likely. The "reference design" and safety case for the Sizewell reactor prepared by the National Nuclear Corporation is now with the Nuclear Installations Inspectorate for first scrutiny, and it includes a major extension of safety features compared with its American model, the Westinghouse reactor under construction at Callaway, Missouri. The independent emergency core cooling systems have been multiplied, the containment vessel extended to a double-wall construction and more concrete introduced into the reactor floor area to protect workers from the relatively high radiation levels of PWRs compared with the old Magnox stations.

Mr Brian George, who was recently appointed by the Central Electricity Generating Board to direct its interests in the PWR project, said on Monday that the board's approach to ordering for Sizewell might differ if it could clearly be seen to be part of a long programme of orders, for then it would be prudent to build up the UK nuclear industry. On the other hand, the first order might be the last. And Mr Reginald Flint, responsible for the reference design at the National Nuclear Corporation, confirmed that all orders for the PWR would have to be placed on a strictly commercial, competitive basis, with no preference for UK suppliers.

Robert Walgate

French hiatus

The French Prime Minister, M. Pierre Mauroy, has confirmed a "freeze" on nuclear power construction in France, and promised a "great debate" in the National Assembly which will question the role and cost of nuclear power in France relative to that of coal and other sources of energy. But the idea of a national referendum on the matter has been pushed to the back of the stage, and the new government has raised no opposition to a major nuclear commitment made by the previous administration.

For in the last few days before the newly elected socialist president took over at the Elysée, the outgoing Prime Minister, M. Raymond Barre, agreed to the doubling of the capacity of the principal French oxide fuel reprocessing plant at Cap de la Hague near Cherbourg, run by the company Cogema.

Cap de la Hague has in recent years been dogged by malfunction and minor accidents. Its existing UP2 400 plant for oxide fuel (as used in pressurized water reactors) was designed to handle 200 tonnes a year; but in five years of operation it has been able to reprocess only 250 tonnes.

M. Barre's decree, if not overturned by the Mitterand government, will allow Cogema to spend 20,000 million francs (£2,000 million) on upgrading UP2 400 to UP2 800 (to handle 800 tonnes a year of spent oxide) and on building UP3, a completely new plant to handle an equal amount of spent fuel from overseas.

The technology of the new plant will be the same as before, but there will be parallel shearing and dissolution sections to give spare capacity if one part of the plant breaks down — small comfort perhaps for Cogema's six customers, Japan, West Germany, The Netherlands, Belgium, Switzerland and Sweden.

Some in France believe that the proposed national debate over nuclear matters could mean that the Cap de la Hague decision could be reopened, and that the decision to reprocess foreign fuel may be reexamined. **Robert Walgate**

Yugoslavian nuclear power

Locals late reacting

The charging of Yugoslavia's first nuclear power station, at Krsko on the River Sava, has provoked a belated burst of opposition. During the construction of the reactor, a 664 MW Westinghouse unit, local feeling was muted. But when charging with uranium fuel was about to begin, the reactor became an issue in the Yugoslav press.

Krsko lies on the border between the republics of Slovenia and Croatia, which jointly built the power station. The commission in charge of reactor safety is likewise an inter-republic body, made up of scientists from the Josef Stefan Physics Institute in Ljubljana and the Rudjer Boskovic Physics Institute in Zagreb.

According to the Yugoslav news agency *Tanjug*, the commission decided in April this year that the trial operation of the station scheduled for July would have to be postponed. The commission is said to have found that measures to protect the environment, the staff and the nearby population had not been implemented, that there was no proper procedure for evacuating the staff in an emergency and no proper fire-fighting equipment. Nor had the problems of storing radioactive waste and treating the coolant water been properly dealt with.

The following day, however, *Tanjug* reported that Dr Maurice Rosen, director of the International Atomic Energy Agency's reactor safety department, had announced that there were no "technological obstacles" to the charging of the station. Nine days later, the Slovenian and Croatian authorities gave permission for charging to go ahead.

According to the Zagreb press, the Yugoslav safety committee tried to insist that certain safety work which Westinghouse and the power station management had planned for after charging, should be done first, but the recommendations were ignored. Significantly, *Vecernji list* of 2 June quotes Dr Marko Brankovic, head of the Croatian Commission for the Human Environment, as saying, "We cannot play around with nuclear reactors as if they were 'political factories'. They must be perfect in every technological respect, or else they must not start". A "political factory", in Yugoslav slang, is a factory erected for prestige reasons, without proper attention to technological details and which proves to be a waste of the usually considerable investment.

Part of the present trouble may be caused by local misunderstanding of the distinction between the charging of a reactor and its start-up. There is also considerable concern about the highly polluted state of the Sava, from Krsko down to its confluence with the Danube. Rumour (downstream in Zagreb) at-

0028-0836/81/240446-02$01.00

tributes the state of the river to the construction of the power station; in fact, the chief offender is a paper-mill which also happens to be sited at Krsko.

Although naturally concerned with safety and environmental questions, the members of the safety committee remain optimistic that the power station will, in the end, be completed, though probably not by the deadline of October/November.

Meanwhile both sides of the belated nuclear debate can find comfort in the recent news that after years of prospecting, a rich uranium deposit (Yugoslavia's second) has been found near Probistip in Macedonia, while in the middle of May a federal Council for Solar Energy was established in Belgrade. **Vera Rich**

London University

More trouble ahead

Some of the smaller colleges of London University may have to close and others amalgamate, but the large colleges with international reputations in many subjects should be kept largely intact. That is the gist of suggestions put forward in the second discussion document of the committee on the reorganization of London University, issued last week.

The problems facing the university have grown since the committee issued its first document last December. Now the university has to contend with an 8 per cent cut in its grant over the next three years in addition to the loss of income from the expected shortfall in overseas students.

The University Grants Committee is expected to tell each university how much money it has for 1981–82 and to allocate student quotas for 1983–84 at the beginning of July. But the London University committee urges that the implications of the expected cut in income are so severe and the timescale so short that discussion of what should be done must start now.

The document says that the university must cater for at least a 10 per cent decrease in the number of students and should reduce the number of buildings it occupies. The suggestion most likely to raise hackles is that Chelsea College, which caters mainly for science and mathematics, be drastically slimmed down or even closed. Chelsea has a good reputation in science education, nursing and the history and philosophy of science, says the committee. But it is not confident that the college's other departments are as good. The good should be sorted out from the bad by peer review, says the committee. Departments surviving the review may be sufficient to form a small college on one site — Chelsea is now widely scattered over several sites. Otherwise they should be incorporated into another college.

The sites that should be kept intact are those supporting colleges with high reputations in a wide range of subjects and, if consistent, those with low costs. Imperial, University and King's colleges and the London School of Economics are likely to emerge largely intact. They may even gain from closer cooperation with some of the smaller colleges — for example, the committee would like to see Queen Elizabeth College, a small but successful institute of science, increase its links with Imperial College. Poor departments in Bedford and Westfield colleges should be weeded out and the two combined on the Bedford site, with the help of money acquired by the sale of Westfield's valuable Hampstead property. **Judy Redfearn**

Research in Poland

Cutting red tape

The Polish production ministries have proved ineffective in coordinating research and development and their role should be phased out, according to Dr Janusz Gorski, Minister of Science, Higher Education and Technology. Writing recently in the weekly *Politkya*, he argued for decentralization and a restoration of decision-making power to the universities and academic institutions.

Academic autonomy, in research as in every other field, has been widely urged in Poland during the last few months by the Academy of Sciences, by the universities, and by the Solidarity "circles" in research and academic institutes. Indeed, Dr Gorski's article was an answer to a swingeing attack, also published in *Polityka*, by Maciej Ilowiecki, one of Poland's leading science writers. According to Ilowiecki, only the restoration of academic autonomy to the scientists could cure the long list of ills of Polish science, including an "absurd" system of financing.

Dr Gorski's article, for the first time, makes some concrete suggestions for improvements. In particular Dr Gorski criticizes the "hierarchy of problems", which has been the basis of Polish science planning since the early 1970s.

According to this scheme, all research funding, with the exception of a small basic research sector, is divided between various "problems" associated with the needs of the national economy. These are classified, in order of importance, as "governmental", "key", "interdepartmental", "departmental" and "branch" problems. Research topics are normally subdivided between a number of institutes and research teams, whose work is coordinated, according to the class of problem concerned, by the Academy of Sciences, and/or the relevant ministries.

Not surprisingly, this scheme has produced a mushrooming bureaucracy, and a scramble among research teams to get their own project as high up the ladder as possible. Now, however, Dr Gorski has proposed eliminating the bottom two levels, "departmental" and "branch", precisely where the role of the production ministries is greatest. In their place, he suggested a system of direct contracts between industry and research institutes, possibly based on competitive tenders.

Whether Dr Gorski's ministry will survive to implement any changes, however, is in doubt. Proposals on rationalizing government spending may lead to the ministry's responsibilities going to enlarged education and industry ministries. In any case some feel that the much-needed reforms cannot be settled while the new legislation on higher education and new statutes for the Academy of Sciences are still only in the discussion stage. **Vera Rich**

Psychiatrist on trial

Dr Anatolii Koryagin, consultant psychiatrist to the Moscow "Working Group to Investigate the Use of Psychiatry for Political Purposes", went on trial in Moscow last week. Until his arrest last February, Dr Koryagin was the last member of the group still at liberty in Moscow — his fellow psychiatrist, Dr Aleksandr Voloshanovich, emigrated in February 1980 and the remaining members of the group are either in prison or labour camp.

Dr Koryagin practised psychiatry in Soviet hospitals for 17 years. Since the working group was set up in 1977, he has examined a number of patients consigned to psychiatric clinics as a result of their conflict with the Soviet State system, and found no sign of disorder.

Although the Soviet authorities have recently abandoned the confinement on phychiatric grounds of well-known dissidents, according to Dr Koryagin, the practice has become widespread for citizens of no great standing who commit some act construed as anti-Soviet. In an article which he sent to *The Lancet* shortly before his arrest, and which was published in April, Dr Koryagin analysed a number of these cases. In many of the cases he cited, the "patients" were simply confined for some weeks or months and then released without receiving any treatment — proof positive, he concludes, that the reason for containing them was isolation, not therapy.

Dr Koryagin was charged with "anti-Soviet agitation" — a far more serious offence than that of slandering the Soviet state, with which several of his fellow members of the commission were charged. In an appeal to foreign colleagues, which reacthed the West a few days before his trial, Dr Koryagin stated that the "transformation of psychiatry into an instrument of politics" was an attempt "drastically to distort the nature of this humane science". While praising the efforts of Western psychiatrists to combat these abuses of their profession, he urged them to be even more decisive in the future.

0028-0836/81/240447-01$01.00

CORRESPONDENCE

Post Office inertia

SIR — It is interesting that a scientific journal, and one as influential as *Nature*, should take up the question of the telecommunications monopoly "deregulation" (*Nature* 26 March, p.279, 23 April, p.619). The importance of this matter is not generally appreciated, and your articles contain some very pertinent comments.

This issue transcends party politics; though no doubt the fate of the British Telecommunications Bill will be influenced by the usual "state ownership versus free enterprise" dogma. The fact is that a country possessing an efficient inexpensive telecommunications system operates with an advantage over one which possesses a relatively expensive inefficient one.

Implementation policy for a universal communications system is important; it would be hard to argue that a remote farmhouse in Cumberland should be without a telephone because resources have been diverted to the profitable business area. But the wealth-producing capacity of the nation mainly resides within manufacturing and service industries, whose ability to create that wealth is substantially affected by the availability of efficient communication services.

It appears that US deregulation policy, falteringly followed since 1968, has not resulted in "cream-skimming" at the expense of services in remote areas. There is an extraordinary range of facilities available in that country which are unavailable here.

I have a list of 60 organizations in the United States offering a range of telephone and data communications equipment; there are probably many more. The number of companies in this country catering for this demand, excluding US subsidiaries, can be counted on the fingers of one hand. The reason is twofold; first, the demand for these devices here is small as yet. Second, the inertia of large bureaucracies encourages a safe policy and the purchase of very long-lasting equipment. The UK telecommunications industry has until very recently been asked to supply out of date equipment to the Post Office. The industry's concentration on the manufacture of this equipment left it in a weak position when tendering for the large international telecommunications contracts which have been placed during the last few years. They were awarded elsewhere.

A.E. CAWKELL

CITECH,
Uxbridge, Mddx, UK

Bovine units

SIR—With reference to Dr John Damuth's letter "Population density and body size in mammels" (*Nature* 23 April, p.699), it has been a general "rule-of-thumb" of the US cattle industry that the annual sustained "carrying capacity" of grazing land is approximately equal to one cow unit per section (square mile) per inch of annual rainfall, with one additional inch of rainfall, with one additional inch of rainfall being assumed for each 1,000 feet of altitude over 5,000 feet; and, a cow unit being defined as a "mother-cow with calf to natural weaning, plus one bull for about each 30 mother-cows."

C.P. ANDERSON

Seattle, Washington, USA

Look harder

SIR — Andrew Brooks in his letter (*Nature* 7 May, p.7) raises a problem which is of wide relevance for the effective application of research and development effort. He is correct in believing that the lack of training in use of information results in much unintentional duplication or misdirection of research. It is particularly galling to us that the area of astronomy in which Mr Brooks presumably undertook his research is covered by the "computerized" *Physics Abstracts* which forms part of the INSPEC database. This database covering physics, electrotechnology and computing is in common with other databases accessible worldwide through a variety of on-line services.

The major difficulty is to make potential users aware of the range of information services and to train them in their use. Inclusion of such training in all undergraduate courses would obviously be a solution, although time for this training is not easily found in heavily-loaded courses. We in INSPEC would be willing to cooperate with universities and colleges to supplement their own efforts at either the undergraduate or postgraduate level.

PETER CLAGUE

INSPEC, Institution of Electrical Engineers,
Hitchin, Herts, UK

SIR — We read with great interest Andrew Brooks' letter (*Nature* 7 May, p.7) concerning the problems of literature searching as part of the PhD research project. For some years we have been tackling this problem at Leicester Polytechnic, where user education is now particularly well-developed. A team of five Academic Librarians has been appointed to train all levels of students in the importance of information and in the techniques involved in its acquisition, storage and retrieval.

We are particularly aware of the daunting size of the task facing postgraduate students in science and technology, when searching the literature as part of their own research. With this in mind we offer courses specifically dealing with the problem at this level, tailored to each postgraduate's own research area.

By means of lectures, demonstrations, practical work and individual tutorials the researchers are introduced to the diverse range of published and unpublished sources in their subject field. The principles and techniques involved in both manual and computerized information searching are explained, and demonstrations are given to show the potentialities of accessing information via on-line bibliographic data-bases. The library offers computerized information retrieval facilities to all research students, many of whom find that a combination of manual and computerized searching provides the best solution. Guidance is also given in related areas such as thesis presentation, report writing techniques and coping with foreign language material.

Similar facilities are provided in several other academic institutions.More widespread encouragement of this trend would go a long way towards solving the problems outlined by Mr Brooks.

RUTH WEBB
CAROLYN HALL

The Library,
Leicester Polytechnic, UK

The fact is . . . .

SIR — Disputes among the supporters of evolution are often only "a matter of presentation". For example, "Evolution is a fact" is a statement I thoroughly dislike, but is understandable to those who have traced the many meanings of "fact". The postulates of a well established theory are sometimes called facts. Mitchell[1] gives the postulates of the chemiosmotic theory and then points out that they are now considered to be facts by most persons. So the postulate-fact confusion arises from the failure to specify one's temporal point of view and adherence to certain views.

The terms "true", "proof", "belief" and "falsification" are often carelessly used when discussing evolution. The first two terms are so final in their meaning that they have mostly disappeared from the discussions of the complex theories. Theories of very limited scope can often be proven true, but the broad, general theories never are. Since "belief" is a favourite term of the creationists and since for them it is something strong and final, scientists could well omit the term, as many have.

Only when one examines the structure of knowledge in evolution can one evaluate "falsification" used in its simple, literal sense. Evolution consists of a system of hundreds of theories[2] which stem from Darwin's two major theories — the kinematic theory of descent with modification and the dynamic theory of natural selection. These theories in Darwinian and neo-Darwinian forms have spawned and are spawning hierarchies of subtheories. Some of the postulates of the major theories and subtheories and some of the the subtheories as a whole may be falsifiable, but usually only new postulates and new subtheories are falsified and discarded. The usual fate of established postulates and theories is not falsification but limitation. Decades ago Delbruck[3] cautioned biologists about the folly of discarding good laws and theories on the grounds that they had limited ranges of applicability.

The testing of a major theory rests with the testing of its subtheories and the fate of the major theory is determined not by a single test, but by an evaluation of the successes and failures in its system of subtheories[4-6]. "Falsification" is not then of much use in evaluating the major theories of evolution. The notion that evolution is not scientific as it is not falsifiable is based on a thin understanding of the structure of major theories, the way they develop and the part being tested.

Since creationists are planning the "ultimate court battle" in which they "will try to establish the 'religious' nature of evolutionary beliefs"[7], it is essential that the scientific community clarify and solidify its views and terminology. We must find and state clearly our fundamental agreements so that our differences in presentation cannot be misconstrued and distorted.

RALPH W. LEWIS

Department of Natural Science,
Michigan State University, USA

1. Mitchell, P. *Science* **206**, 1148–1159 (1979).
2. Lewis, R.W. *Perspectives biol. Med.* **23**, 551–572 (1980).
3. Delbruck, M. *Trans. Conn. Acad. Arts Sci.* **38**, *173 (1949).*
4. Bunge, M. *Appl. math. Modeling* **2**, 201–204 (1978).
5. Caplan, A. *Erkenntis* **13**, 261–278 (1978).
6. Tuomi, J. & Haukioja, E. *Savonia* **3**, 1–8 (1979).
7. Broad, W.J. *Science* 1331–1332 (1981).

0028-0836/81/240448-01$01.00

Overfishing: is there a solution?

from A.J. Southward

THE late Michael Graham, sometime Director of United Kingdom Fishery Research, always maintained that "fisheries that are unlimited become unprofitable . . . or inefficient"[1]. Nevertheless, argument continues over the extent to which observed changes in fish populations are the result of overfishing or of natural causes, including climatic change[2-4].

A good example of such disagreement can be seen in the recent history of the formerly profitable herring fishery in the southern North Sea. Some scientists believed that the decline after 1955 was in part a natural change, others that it was due to overfishing of the adult shoals, and yet others that the industrial fishery for juveniles was responsible[5]. As a result of the failure to agree about the causes, conservation measures were delayed, and the fishery collapsed. Even today some scientists are unconvinced that fishing can reduce stocks to below the point of recovery and cause changes in the whole ecosystem. For this reason two recent communications to *Nature* are significant: one pointing towards the causes of a severe decline in certain species (K. Brander, vol.290, p.48), the other showing massive changes in the ecosystem that may have resulted from removal of predators by overfishing (K. Sherman *et al.*, this issue, p.486).

In the case of the common skate, *Raia batis*, Brander has developed the classic method of comparing 'recruitment' — essentially the sum of breeding intensity and survival of the young up to the size when they are taken by fishing gear — with the subsequent mortality each year afterwards due to natural causes and fishing. Skate lay only a few large eggs each year, and although these are well protected by a horny case, the young that hatch from them are at once large enough to be captured by the meshes of standard trawls. Skate also take a long time to reach maturity — as much as 11 years. Thus even a moderate level of fishing is enough to lead to virtual extinction of the population. This explains the almost total disappearance of skate from the Irish Sea during the past 20 years, and may explain a much earlier decline in the western English Channel[6]. Other species of *Raia* are nearly as vulnerable as skate to a moderate level of fishing, and another elasmobranch fish that produces only a few young per year, the spur dog, *Squalus acanthias*, may be equally endangered[7]. This species virtually disappeared from the western Channel between 1920 and 1950[8], and more recently has declined in the northern North Sea[9].

In contrast to the elasmobranchs, most teleost fish produce immense numbers of eggs, and it used to be argued that the effects of fishing on these species must be very small compared with the great natural mortality during the early stages of their life history[10]. This argument was probably not true even a hundred years ago[1], and today's technology, especially the use of sophisticated sonar to locate the shoals and large purse-seines to trap them, can produce fishing mortalities up to ten times greater than the natural mortalities[11]. Removal of a species on a scale like this must surely influence the rest of the ecosystem, for example by allowing increases in abundance of a competing species or of organisms that were eaten by the fish. Some scientists believe this happened in the 1970s in the northern North Sea, where, after the decline of herring and mackerel, there were increased catches of 'trash' species such as small gadoids, sand eels and sprat[12,13]. This 'gadoid outburst' has also been attributed to natural causes, such as an improved 'match' between the spawning season of the fish and peak production of the plankton on which the young feed in the spring, possibly related to a trend in climate[2,14].

The detection of a similar outburst across the Atlantic by Sherman and colleagues is therefore significant. These workers report increases in abundance of young sand eels, during the past 6 years, over a wide area of sea off the north-east coast of the United States, from George's Bank to Cape Hatteras. Although there has been variation between years in the position of maximum abundance, all three major fishing regions surveyed show increases of more than an order of magnitude in the number of sand eels. The stocks of herring and mackerel in the same regions appear to have declined greatly since 1972, and it is difficult to avoid the conclusion that removal of some of its major predators has allowed the 'opportunistic' sand eel to increase.

However, this increase might also be attributed to natural changes. Good correlations can be shown between sea-temperature trends and fluctuations of several commercial fishes in the Gulf of Maine and off Nova Scotia[15,16]. Similar correlations can be shown for UK waters, but in each case the variations in catches are small and are correlated with quite small changes in sea temperature.

With the gadoid outburst in the North Sea, and the increase in sand eels in the North-West Atlantic, we are faced with much greater changes in abundance. A comparable change in the ecosystem, with differences of one or two orders of magnitude in the numbers of the young of demersal fish, took place in the western Channel: first a decrease when herring gave way to pilchard as the dominant pelagic fish, then an increase when the pilchard declined and mackerel moved in[17]. This change in the English Channel has always been accepted as natural[2], probably related to the effect of climatic fluctuation on the position of a boundary between different ecosystems[17]. The events in the North Sea and the North-West Atlantic are difficult to relate to changes in climate, because they have occurred over a wide range of latitude and there is little sign of shifts in geographical distribution of key species or of organisms from other habitats, such as were reported from the English Channel[4,17]. It is therefore simpler to accept, for the moment, the hypothesis that the increases in 'trash' species are due to removal of herring and mackerel. Such an hypothesis should be easy to test, by

---

1. Graham, M. *The Fish Gate* (Faber, London, 1943).
2. Cushing, D.H. & Dickson, R.R. *Adv. mar. Biol.* **14**, 1 (1976).
3. Cushing, D.H. *Phil. Trans R. Soc.* **B286**, 597 (1979).
4. Southward, A.J., Butler, E.I. & Pennycuick, L. *Nature* **253**, 714 (1975).
5. Burd, A.C. in *Sea Fisheries Research* (ed. Harden Jones, F.R.) 167 (Elek, London, 1976).
6. Unpublished records, Marine Biological Association, Plymouth.
7. Holden, M.J. *Fish. Invest., Lond.* **II**, 25 (8), 1 (1968).
8. Southward, A.J. *J. mar. biol. Ass. UK* **43**, 1 (1963).
9. Holden, M.J. in *Sea Fisheries Research* (ed. Harden Jones, F.R.) 117 (Elek, London, 1976).
10. Huxley, T.H. *International Fisheries Exhibition, 1883* **IV** (1), 2 (1884).
11. Burd, A.C. *Rapp. P.-v. Réun. Cons. perm. int. Explor. Mer* **172**, 137 (1978).
12. Andersen, K.P. & Ursin, E. *Rapp. P.-v. Réun. Cons. perm. int. Explor. Mer* **172**, 286 (1978).
13. Andersen, K.P. & Ursin, E. *Meddr. Danm. Fisk. -og Havunders.* (N.S.) **7**, 319 (1977).
14. Cushing, D.H. *J. Cons. Explor. Mer* **39**, 70 (1980).
15. Dow, R.L. *J. Cons. Explor. Mer* **37**, 274 (1977).
16. Sutcliffe, W.H.J., Drinkwater, K. & Muir, B.S. *J. Fish. Res. Bd Can.* **34**, 19 (1977).
17. Southward, A.J. *Nature* **285**, 361 (1980).

0028-0836/81/240449-02$01.00

stopping or reducing the fishing of the herring or mackerel, whereupon the numbers of 'trash' species should fall; but whether effective international action can be achieved to sustain such a test is another matter.

No 'easy' solution is available for the elasmobranch fishes, which are endangered by such a low level of fishing that only a complete ban on all fishing over a wide area would allow recovery. As Brander says, we have to face the prospect that skate, and similarly sensitive species, will be fished out (or become extinct) as a consequence of the exploitation of other fish. □

A.J. Southward is at the Plymouth Laboratory of the Marine Biological Association of the United Kingdom.

Collagen evolution

from Ellen Solomon and Kathryn S.E. Cheah

COLLAGEN has a long evolutionary history, and has been found in all invertebrate and vertebrate phyla studied, including such primitive invertebrates as fresh-water sponges and sea anemone. Considerable advances have been made in the understanding of the biochemistry of collagen, and its precursor procollagen, with respect to molecular structure, assembly and post-translational processing[1-3]. Until recently, however, attempts to analyse the genetic origins of this protein have had to rely on amino acid sequence data which, because of the size, complexity and difficulty of isolation of the collagen molecules, have been difficult to obtain. Recombinant DNA technology and DNA sequencing methods now provide an alternative means of analysing the origins of the collagen genes. During the past year the structure of one such gene, $\alpha2$(I), has been determined in some detail, and this in turn has provided a substantial clue as to its evolution[4,5].

Basic to all collagenous proteins is the amino acid tripeptide Gly-X-Y, which is repeated approximately 330 times in the most abundant and best studied, Type I collagen, about 1,600 times in annelid cuticle collagen[6], and approximately 29 times in the chains of C1q[7]. These molecules are all triple helical structures, comprised either of three identical chains, or two different chains, in the ratio of 2:1. Recently a new dimension has been added to the biology of connective tissues with the discovery that, in addition to being a very ancient molecule, collagen is also exceedingly complicated genetically: it exists, at least in vertebrates, in at least five distinct types, for which a minimum of nine genes must exist to code for their constituent chains[8,9]. If proteins with collagenous domains such as C1q[7] and acetylcholinesterase[10] are added to this list, the number of genes in the family increases to at least 15 or 16.

The evolution of these genes is interesting from several points of view. Because of the periodic nature of the Gly-X-Y repeat in these proteins, and the general conservation of sequences, it has been assumed that the collagen genes are all related evolutionarily through duplication and divergence of an ancient collagen gene. However, although the fundamental unit which expanded was clearly the Gly-X-Y tripeptide, the question has been raised as to whether there was an initial expansion of this triplet to a small primordial gene. This small gene could then have duplicated to give a long chain, which itself duplicated and diverged, or could have been the basis for the independent generation of several long chains.

Attempts to answer such questions and to evaluate the evolutionary relationships between the collagen chains have relied on statistical analysis of amino acid sequence data. Most of the available data are for the $\alpha1$(I) and $\alpha2$(I) chains, with some data for $\alpha1$(II) and $\alpha1$(III)[3]. As much of the sequence data is incomplete and could be non-representative of the entire chain, the conclusions about evolutionary distances are quite provisional. Nonetheless, calculations[11] indicate that the $\alpha1$(I) and $\alpha2$(I) chains diverged about 7×10^8 years ago, about the time of the separation of vertebrates and invertebrates. The $\alpha1$(I) chain has evolved slowly, at about the rate of cytochrome *c*, while the $\alpha2$(I) chain has evolved more quickly, at about the rate of haemoglobin α and β. On the basis of sequence comparisons of four α chains, an evolutionary tree has been proposed, in which successive gene duplications resulted in separate genes coding for $\alpha1$(III), $\alpha2$(I), $\alpha1$(II) and $\alpha1$(I)[12], in that order. Although some of these calculations may be provisional, the data clearly show that the conservation of sequence of each of the chains is far greater for the same chain across species (variation in about 5–10 per cent of non-glycine residues) than it is between different chains within a species (variation in about 30–60 per cent of non-glycine residues).

A model of a primordial collagen gene which duplicates to form one chain would predict that internal homologies would be found within that chain, provided, of course, that the sequence were sufficiently conserved for these to be detected. McLachlen[13], using a comparison matrix on the amino acid sequence of the $\alpha1$(I) chain, has detected significant regularities occurring at a distance related to the stagger between adjacent triple helical molecules in the collagen fibre. These occur at a distance, *D*, of 234 amino acids, or 78 triplets. He suggests that if gene duplication has occurred, the repeated domain is likely to be of length *D*, or a simple fraction of *D*. More recently, Hofmann, Fietzek and Kuhn[14], analysing the $\alpha1$(I), $\alpha2$(I) and $\alpha1$(III) chains, also found the *D* repeat, but in addition found smaller repeats of *D*/3, *D*/6, *D*/11 and, perhaps more significantly, *D*/13 (18 amino acids). Also emerging from these data is the fact that the homologies within one repeat are more conserved across species than from one repeat to another within a chain. The most likely explanation for this is early intrachain duplication with subsequent duplication of the whole chain.

Recently, the application of recombinant DNA technology to the analysis of the collagen $\alpha2$(I) gene, and the subsequent determination of DNA sequences, have yielded more definitive predictions of the size of the hypothetical primordial collagen gene. The isolation and characterization of genomic clones covering approximately half the sheep $\alpha2$(I)[15,16] and the whole of the chick $\alpha2$(I) gene[17,18,5,6] have revealed that this gene is extremely large and has the greatest number of intervening sequences of any gene studied to date[19]. Five kilobases (kb) of coding sequence are distributed over at least 38 kb of DNA containing 49 introns varying in size from less than 100 base pairs (bp) to more than 3,000 bp. Introns are largest and exons smallest within the region coding the triple helical portion of the polypeptide. More strikingly, DNA sequence analysis of some of the exons within this portion of the chicken gene reveal that these exons always end at a position corresponding to the Y position of the Gly-X-Y triplet and never in the middle of a triplet[5,6].

Yamada and colleagues[5] have sequenced eight exons of the chick $\alpha2$(I) gene and find that seven of these are 54 bp (18 amino acids) long, while the eighth is 99 (54 + 45) bp. Five additional exons have now been sequenced by the same group — three of 54 bp and two of 108 (2×54) bp (B. de Crombrugghe, personal communication). These authors therefore propose that the first collagen gene arose by amplification of a single coding unit of 54 bp. They predict that all collagen genes, whether in sponges, nematodes, or man, will be found to have exons of this length. In addition, this small primordial gene may have functioned in the primitive cell, and small collagenous regions, such as those found in C1q and acetylcholinesterase, could be remnants of such a function. The occurrence of an exon of 99 bp can be explained by various recombinational and

Ellen Solomon and Kathryn S.E. Cheah are at the Imperial Cancer Research Fund Laboratories, Lincoln's Inn Fields, London.

0028-0836/81/240450-02$01.00

mutational events, all of which must involve the deletion of 9 bp (corresponding to one Gly-X-Y triplet). Yamada and co-workers predict that for any collagen gene for which the protein helicity is conserved, all insertions or deletions in the helical coding regions will be 9 bp or multiples of this length.

Wozney and colleagues[6] have arrived at a similar conclusion based on the analysis of seven exons derived also from the chick α2(I) gene. Additional data from this laboratory bring the total number of exons which they have sequenced to fourteen (H. Boedtker, personal communication) of which seven are 54 bp, two are 108, two are 45 and three are 99. To accomodate the exons of 45 and 99 bp, these authors suggest that, during evolution, duplication of the 54-bp gene was accompanied by additions and deletions of 9 bp.

One of the exons studied by Wozney *et al.* contained the coding sequences covering 45 bp at the end of the triple helical region of the pro α2(I) chain as well as 204 bp of the adjacent non-triple helical carboxy-terminal domain. Included in this non-triple helical region is the procollagen peptidase cleavage site. This is somewhat surprising since according to the concept of introns separating functional domains, one might have expected to find an intron separating sequences specifying helical and globular regions of the procollagen molecule, and perhaps also at the peptidase cleavage site. Any hypothesis involving the duplication of a 54-bp primordial collagen sequence to give rise to a procollagen gene must therefore also explain how this sort of exon could have arisen. Loss of introns is the most likely possibility.

Although the hypothesis of evolution of the collagen genes by duplication of a small exon of 54 bp is attractive, it must be remembered that so far it is based solely on the α2(I) gene with the majority of exons not yet analysed. Data on some of the other genes are eagerly awaited. □

1. Prockop, D.J., Kivirikko, K.I., Tuderman, L. & Guzman, N.A. *New Engl. J. Med.* **301**, 13 (1979).
2. Heathcote, J.G. & Grant, M.E. in *The Enzymology of Post-Transcriptional Modification of Proteins,* 457 (Academic, London, 1980).
3. Bornstein, P. & Traub, W. *The Proteins* **4**, 411 (1979).
4. Yamada, Y. *et al. Cell* **22**, 887 (1980).
5. Wozney, J., Hanahan, D., Morimoto, R., Boedtker, H. & Doty, P. *Proc. natn. Acad. Sci. U.S.A.* **78**, 712 (1981).
6. Tanzer, M.L. *Trends biochem. Sci.* **3**, 15 (1978).
7. Reid, K.B.M. & Porter, R.R. *Biochem. J.* **155**, 19 (1976).
8. Bornstein, P. & Sage, H. *A. Rev. Biochem.* **49**, 957 (1980).
9. Gay, S., Gay, R.E. & Miller, E.J. *Arthritis Rheumat.* (1980).
10. Rosenberry, T.L. & Richardson, J.M. *Biochemistry* **16**, 2550 (1977).
11. Mathews, M. in *Biology of Collagen,* 173 (Academic, London, 1980).
12. Grant, M.E. & Jackson, D.S. *Essays Biochem.* **12**, 77 (1976).
13. McLachlen, A.D. *J. molec. Biol.* **107**, 159 (1976).
14. Hofmann, H., Fietzek, P.P. & Kuhn, K. *J. molec. Biol.* **141**, 293 (1980).
15. Boyd, C.D. *et al. J. biol. Chem.* **255**, 3212 (1980).
16. Schafer, M.P., Boyd, C.D., Tolstoshev, P. & Crystal, R.G. *Nucleic Acids Res.* **8**, 2241 (1980).
17. Vogeli, G. *et al. Nucleic Acids Res.* **8**, 1823 (1980).
18. Ohkubo, H. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 7059 (1980).
19. Cheah, K.S.E. & Grant, M.E. in *Collagen in Health and Disease* (Churchill Livingstone, Edinburgh, in the press).

Polymers by synchrotron radiation

from Paul Calvert

SMALL-ANGLE neutron scattering has made a dramatic impact in polymer physics over the past few years. It is therefore interesting to speculate whether anything similar might come from synchrotron radiation, especially in view of the fact that the synchrotron radiation source (SRS) at the SRC Daresbury Laboratory was commissioned last June.

In brief, the SRS provides an intense, continuous source of radiation from the hard X-ray region out to microwaves. The radiation is predominantly polarized and pulsed in about 120-ps pulses at 2-ns intervals. What can actually be done at Daresbury will depend on the installation of suitable monochromators, cameras and detectors. In the X-ray region the available intensities are three to five orders of magnitude greater than from conventional rotating anode sources, permitting exposures of a few seconds in wide- and small-angle diffraction so that time-dependent structural changes can readily be studied. In the UV, visible and IR, the intensities available can be equalled by arc lamps or lasers though the continuous spectrum of the SRS may be an advantage. In the long-wavelength range, synchrotron radiation again becomes more intense than conventional sources.

Three reports published recently on the use of synchrotron radiation for X-ray studies of polymers all concentrate on the measurement of changes during phase transformations. Forgacs (Budapest) together with co-workers at Novosibirsk report the use of the Vepp-3 storage ring at the Institute of Nuclear Physics in Novosibirsk to study melting crystallization and drawing in polyethylene and polypropylene (*J. Polymer Sci.* **18**; 2155, 1980). A position-sensitive counter was used to record the scattering from 7 to 35°, 2θ, into 1,000 channels. The authors claim to get recognizable curves within 1 second but generally used counting times of 30-60 seconds. They present data on high- and low-density polyethylene and polypropylene during heating from room temperature through the melting range. Time dependence is not reported, possibly because none was seen. The time taken for the sample to reach thermal equilibrium at each temperature was several minutes which may exceed the time required for changes in the structure. Crystallization of polypropylene could be followed as a function of time at 120°C although again the process was half over before thermal equilibrium was reached. They were also able to follow the relaxation of a sample after rapid uniaxial stretching.

Elsner, Zachmann and Milch used the DORIS storage ring at DESY in Hamburg to study crystallization of oriented polyethylene terephthalate on heating from the glassy state (*Makromolec. Chemie* **182**; 657, 1981). The complete small-angle X-ray pattern was recorded over an area of diameter 4 cm using an X ray-sensitive vidicon tube. Exposure time was 10 seconds. The long period, which is a measure of crystal thickness, was found to decrease during crystallization, in agreement with previous studies made on partly crystallized and quenched samples. Also at DESY, Koch and co-workers (*Polymer Bull.* **1**; 709, 1979) measured crystallization in stretched polyisobutylene rubber. Samples were extended in 0.5 second and data then collected over a series of intervals of 5–30 seconds. The crystallization was shown to follow the Avrami equation with an exponent of one. When the sample was allowed to retract the crystallinity disappeared within 5 seconds.

Shorter exposures than this have been achieved with conventional sources. Using a rotating anode tube Stein and co-workers (*J. Polymer Sci.* **46**; 15, 1974) were able to obtain rather fuzzy diffraction patterns from polyethylene in 1/60th second and observed the development of a single peak in stretched polyethylene with a resolution of about 0.025 second. Clearly the synchrotron source has the potential to improve greatly on this if the sample can be stretched rapidly enough. Very little is known about the melting kinetics of polymers and dynamic X-ray studies using the SRS could be very useful, but this requires sample holders capable of reaching thermal equilibrium in a few seconds. On the other hand, a suitable choice of samples and conditions will give crystallization rates which can be easily followed. However, because the process is intrinsically heterogeneous, we really need to know what happens in a very small volume of material as crystals grow through it; this means a beam size of the order of 1 μm.

These applications only skim the surface of the potential uses of the SRS (for a review, *see Synchrotron Radiation Research* eds H. Winich and S. Doniach, Plenum, 1980). Extended X-ray absorption fine structure (EXAFS) has proved very informative about metal ions in enzymes and has great promise for the study of surfaces and of catalysts. X-ray topography and soft X-ray microscopy using the SRS also look interesting. Synchrotron radiation is a very powerful spectroscopic tool; the challenge is to think of how to use it. □

Paul Calvert is at the School of Molecular Sciences, University of Sussex.

0028-0836/81/240451-02$01.00

Atom-specific bond breakage

from Michael L. Knotek

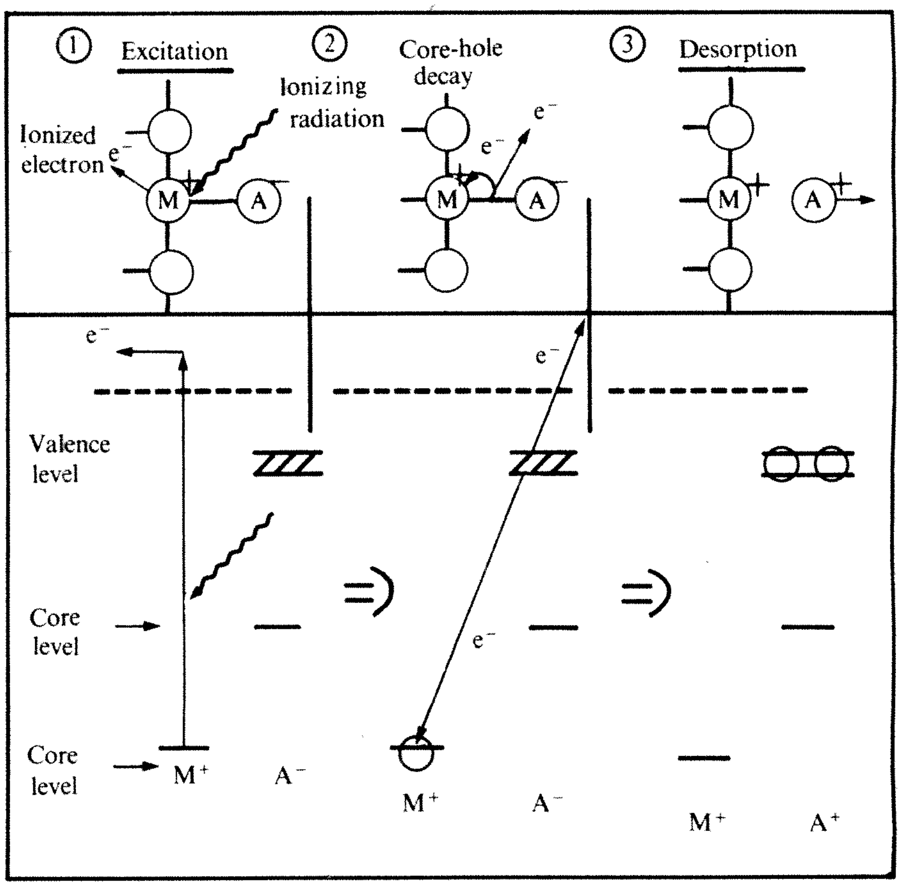


Fig. 1 Ions are emitted from surfaces (or bonds broken in a molecule or solid) in a three-step process. In an example of ionic bonding (M^+A^-): 1, Ionizing radiation excites an electron from a core level in a bonding site atom Ⓜ$^+$. 2, The resulting 'core-hole' decays by absorbing a valence electron and emitting one (or two) more. 3, The resulting multiple valence hole final state is highly repulsive and the resultant adsorbate, Ⓐ$^+$, desorbs.

A recent series of discoveries in the field of electron- and photon-stimulated desorption (ESD and PSD) promise to deepen understanding of surface and molecular chemistry and the interaction of ionizing radiation with matter. Those studying ESD and PSD have long recognized that a low-energy ionizing particle (an electron or photon with $0 < \text{energy} < 10^4$ eV) causes bond dissociation by electronic excitation[1,2]. Now it has been found that bonds to a specific surface atom can be broken by exciting an inner shell or 'core' electron on that atom[3–8]. Furthermore, on an ordered surface the emitted (desorbed) ions are seen to emerge in well defined beams which reflect the geometry of the bonding configuration[7,9]. Both of these features are due to highly localized energy deposition and bond breakage phenomena[3]. By using a simple electron source or the intense, tunable light available from a synchrotron[4,6,7,10], the electronic and geometric structure of specific species bonded to specific surface atoms can be studied. The techniques and concepts derived in studying the surface bond and its dissociation by ionizing radiation will be widely applicable to a variety of important chemical systems.

Stimulated desorption experiments have been used to characterize some of the fundamental properties of a bond by studying its dissociation with ionizing radiation. In a typical experiment the energy of an ionizing particle is increased until threshold for desorption of ions is noted. Significantly, the desorption threshold of a given species coincides with the excitation of one of its own core electrons, or a core electron of the atom to which it is bonded[3–8], as was first demonstrated for the desorption of positive oxygen ions (O^+) from TiO_2. The O^+ desorbed when a 'core-hole' was created (that is, a core-level ionized) on the Ti site to which it was bonded[3]. It was subsequently demonstrated that the core-hole decays by having a valence or bonding electron fall into the vacant core-hole. One or more valence electrons will also be ejected to carry off the energy released in the decay (so-called Auger electrons). For an ionic solid the decay of a core-hole on a metal atom (cation) can remove enough charge from its neighbouring anion to cause the latter to change from a negative to a positive ion and hence be in an extremely repulsive state from which it will desorb as a positive ion (Fig. 1). The mechanism is not as simple for covalently bonded solids, but the removal of two or more electrons from the bond can often cause it to break with the emission of ions — a so-called Coulomb explosion[5].

Two crucial features of this kind of bond dissociation are that the excitation of a core-hole stores an enormous amount of highly localized energy in an atom and that the decay of this hole produces multiple valence-hole states localized on the order of atomic dimensions. The concept of intrinsic localization of the Auger final state was explored for a range of chemical systems in relation to the Auger process and now to desorption and dissociation[11–15]. Two or more valence holes created on an atom or in a specific bond strongly repel each other. Contrary to simple intuition, the more strongly the holes repel one another, the longer they are locked together[11,12]. Since this unstable configuration cannot be relieved through electronic processes in a time that is short compared with the characteristic desorption time, localized bond breakage occurs when the valence holes are properly arranged and the electronic state has a sufficient lifetime.

These results clearly demonstrate that the bonding site for a given chemical species can be determined by measuring the desorption threshold and matching that energy with tabulated core-level energies of

---

Michael L. Knotek is in the Surface Physics Division, Sandia National Laboratories, Albuquerque, New Mexico.

1. Menzel, D. & Gomer, R. *J. chem. Phys.* **41**, 3311 (1964).
2. Redhead, P.A. *Can. J. Phys.* **42**, 886 (1964).
3. Knotek, M.L. & Feibelman, P.J. *Phys. Rev. Lett.* **40**, 964 (1978); *Surf. Sci.* **90**, 78 (1979). Feibelman, P.J. & Knotek, M.L. *Phys. Rev.* **B18**, 6531 (1978). Knotek, M.L. *Surf. Sci.* **91**, L17 (1980); **101**, 334 (1980).
4. Knotek, M.L., Jones, V.O. & Rehn, V. *Phys. Rev. Lett.* **43**, 300 (1979).
5. Franchy, R. & Menzel, D. *Phys. Rev. Lett.* **43**, 865 (1979).
6. Woodruff, D.P. *et al. Phys. Rev.* **B21**, 5642 (1980).
7. Madey, T.E., Stockbauer, R., Van der Veem, J.F. & Eastman, D.E. *Phys. Rev. Lett.* **45**, 187 (1980).
8. Knotek, M.L., Jones, V.O. & Rehn, V. *Surf. Sci.* **102**, 566 (1981).
9. Czyzewski, J.J., Madey, T.E. & Yates, J.T. Jr *Phys. Rev. Lett.* **32**, 777 (1974). Madey, T.E., Czyzewski, J.J. & Yates, J.T. Jr *Surf. Sci.* **49**, 465 (1975); **57**, 580 (1976).
10. Jaeger, R. *et al. Phys. Rev. Lett.* **45**, 1870 (1980).
11. Sawatsky, G.A. *Phys. Rev. Lett.* **39**, 504 (1977).
12. Cini, M. *Sol. St Comm.* **24**, 681 (1977); **20**, 605 (1976).
13. Rye, R.R., Jennison, D.R. & Houston, J.E. *J. chem. Phys.* **73**, 4867 (1980). Jennison, D.R. & Kelber, J.A. *Proc. 27th natn. Symp. Am. Vacuum Soc.*
14. Dunlap, B.I., Hutson, F.L. & Ramaker, D.E. *Proc. 27th natn. Symp. Am. Vacuum Soc.* White, C.T., Ramaker, D. & Murday, J.S. *Bull. Am. phys. Soc.* **26**, 396 (1981).
15. Feibelman, P.J. *Surf. Sci.* **102**, L51 (1981).

0028-0836/81/240452-02$01.00

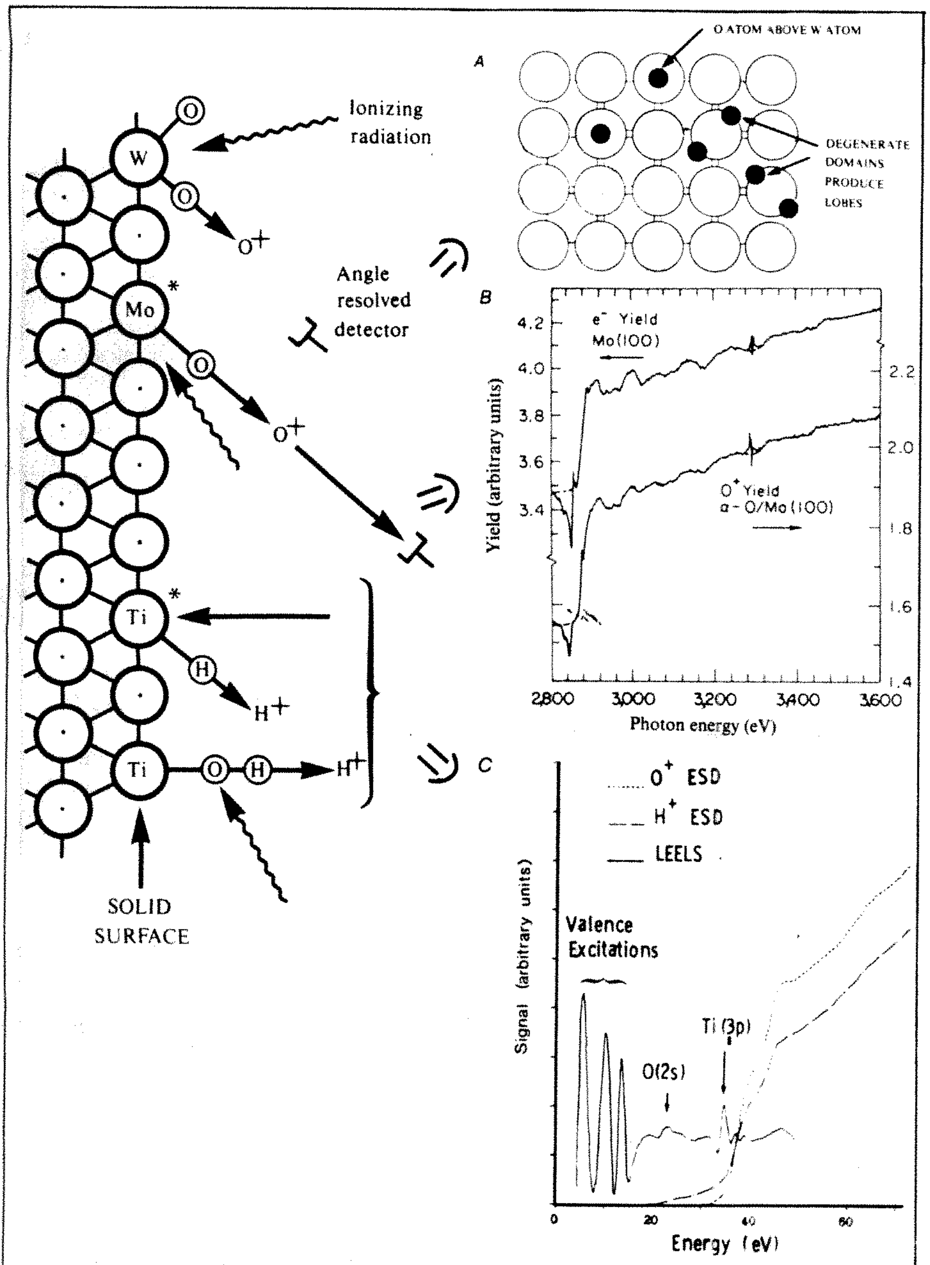


Fig.2 Excitation of specific surface atoms breaks bonds only to those atoms. The ions are emitted in well defined beams which can be observed by imaging devices to give directly an image from which the bonding structure as in *A* can be derived. Structure observed in Xray-excited desorption can be analysed to determine the geometric structure in the vicinity of the excited atom (*B*). For an adsorbate such as C-O on a surface, excitation of the carbon (ls) level gives both Co^+ and O^+ in desorption while excitation of the oxygen (ls) leves gives O^+ (*C*). If a given species (say hydrogen) is bonded to more than one kind of surface atom, then threshold for desorption will be observed at both bonding site atom core level ionization potentials as in the example of hydrogen on TiO_2.

the elements. Each atom has a distinct set of core-level energies which makes the task straightforward and unambiguous. Since apparently only first nearest-neighbour bonds are broken, this technique can determine the bonding sites of all the desorbing species. One problem, however, is that some species do not desorb, desorb weakly, or desorb only as neutrals.

Even more fruitful is the study of the behaviour of the ion yield of a particular species as a function of energy above the excitation threshold. These thresholds are not simply abrupt changes in yield, but contain complex spectral structure within roughly 20–30 eV of the edge (so-called near edge structure) and extending hundreds of electron volts above the edge (extend fine structure). These structures arise because the desorption of ions will simply be proportional to the cross-section for core-hole creation on the adsorption site or adsorbate atom. The spectrum reflects the X-ray absorption spectrum of the excited atom and yields a detailed picture of its electronic and geometric environment. This specificity allows examination of a surface with a variety of adsorbates and bonding sites by extracting the excitation spectrum of each site/adsorbate combination independently of the others.

The correlation of core-hole excitation thresholds with stimulated desorption yield thresholds has been made for metals, semiconductors and insulators[3-10]. The extraction of information from the structure in the desorption yield spectrum is in an early stage of development, but has been successfully applied to a PSD Surface Extended X-ray Absorption Fine Structure (PSD/SEXAFS) study of oxygen on molybdenum[10]. The radial distribution function (r.d.f) of the Mo surface atoms originally bonded to desorbed O^+ ions was determined from modulations in the yield spectrum above the Mo (2s) absorption edge. The bonding site molybdenum r.d.f. (basically a measure of the Mo–Mo distance) was found to remain essentially unchanged upon adsorption of oxygen on the surface — a rather surprising result. The desorbed oxygen apparently comes from MoO_3-like species formed during oxidation. SEXAFS structures were also seen in the desorption of ions from BeO, Al_2O_3 and SiO_2 surfaces[8].

Distinct beams of ions in stimulated desorption were first observed in ESD experiments[9] (ESDIAD for ESD Ion Angular Distributions), and more recently in PSD[7]. ESDIAD patterns reflect the site symmetry and bond angles of adsorbed species on several metal surfaces, and are also sensitive to surface details such as adsorption at step sites, and thermal vibrations of the adsorbate. Significantly, these measurements define the local geometry of the surface–adsorbate bond in a statistical sample of the surface and do not require long-range order to determine structure. ESD/PSD thus has the unique capability to determine structure both from angle-resolved techniques and from analysis of spectral features. The former gives the adsorbate structure (bond angle and symmetry) while the latter characterizes the environment of either the adsorbate or bonding site atom.

Although research has been confined to systems of traditional interest to surface scientists, these studies can easily be extended to more complex structures including organometallics — it may be possible to probe the central metal atom in porphyrins, or perhaps cell membranes where the environment of a specific metal atom can be monitored. This experimental approach is extremely sensitive to hydrogen (which desorbs most readily of all species) and to the details of the electronic and structural nature of hydrogen bonding; so its application to problems involving hydrogen chemistry will undoubtedly be important.

New concepts of desorption and dissociation are being explored and will reveal new features of desorption phenomena and deepen understanding of surface chemical processes. Probably the most important result of this research is a new perspective for surface bonding. This will lead to new ideas not only regarding surfaces, but also radiation-induced chemistry and damage. ☐

0028-0836/81/240453-02$01.00

Neuropeptides: do they integrate body and brain?

from Susan D. Iversen

DE WIED AND BOHUS were the first to demonstrate that an extract of the posterior pituitary delayed the extinction of active shock avoidance behaviour in the rat[1]. Subsequent experiments indicated that the active substance was vasopressin, and synthetic Lys^8-vasopressin given in a single dose (3.0 μg) systemically caused a long-lasting delay in extinction[2]. These results have led to the hypothesis that vasopressin modulates processes underlying memory consolidation. If this claim were to be substantiated, vasopressin might have useful therapeutic effects in alleviating memory disorders in man, such as those seen in Korsakoff's psychosis and Alzheimer's disease.

An important question concerns the site at which vasopressin acts to produce its effects on memory. A potent analogue, Arg^8-vasopressin, has a significant effect on extinction behaviour when administered intraventricularly in doses as low as 0.025 ng. When the dose–response data are compared for the subcutaneous and intracerebral routes of administration, it is clear that much smaller doses are effective when injected directly into brain[3]. Thus, de Wied and his colleagues favour a central nervous system (CNS) site of action. Furthermore, using immunohistochemical staining techniques it has been demonstrated that vasopressin-containing neurones of the magnocellular hypothalamic nuclei project widely to innervate various autonomic centres in the brain stem and spinal cord, and vasopressin fibres are also found in limbic septum[4]. The enhanced memory performance after local injections of arginine vasopressin (AVP) into these limbic areas, has been correlated with changes in CNS noradrenaline function[5]. A number of intriguing facts support the hypothesis that an interaction between vasopressin and limbic forebrain noradrenaline is important for cognitive functions related to memory. Studies on human brain have recently revealed that the locus coeruleus contains the highest level of vasopressin outside the hypothalamus[6]. The locus coeruleus is the origin of the noradrenaline innervation of hippocampus, and lesions to this system result in learning impairments in animals when trained responses must be modified[7]. Degeneration of the locus coeruleus has been observed in the brains of patients dying with Alzheimer's disease[8] and degeneration of the dorsal noradrenaline pathway has been implicated in Korsakoff's disease[9], an amnesic syndrome said to be improved by vasopressin treatment[10].

The question of a central site of action of AVP is raised again in this issue of *Nature* (p.491). Le Moal and colleagues have confirmed the effects of systemic AVP on extinction behaviour in the rat, and report that a novel synthetic antagonist of AVP can abolish this behavioural effect. Within the same dose range, they also demonstrate an effect of AVP on blood pressure, which was reversed by the antagonist at doses effective against the behavioural response. The authors suggest that there is a causal relationship between the effect of AVP on blood pressure and the prolongation of extinction behaviour, and that the visceral arousal induced by AVP and reflected in the blood pressure change induces, via visceral afferents to the CNS, cerebral arousal "sufficient to alert the rat and to cause the perseverance of ongoing or escape behaviours".

They suggest that the reported effects of AVP on brain may also be due to peripherally triggered activation of visceral integrative centres. In the experiments where AVP has been injected into the ventricles or limbic sites, it is possible that the circumventricular organ systems may be the site of action rather than structures within the brain itself. It seems premature to reject one or other of these hypotheses. For example, the vasopressin analogue Desglycinamide lysine vasopressin, which is practically devoid of classical endocrine activity, has an effect on extinction of avoidance behaviour[11]. It would seem that a presser response to vasopressin is sufficient but not necessary for the effect on behaviour.

A more important implication of the paper is that vasopressin and other neuropeptides may act at both peripheral and central nervous system sites in an integrated fashion. Vasopressin released in brain would appear to act on the noradrenergic pathway arising in locus coeruleus. It is interesting that visceral afferent information reaches brain stem areas (including locus coeruleus) rich in vasopressin innervation. In certain behavioural situations visceral and central arousal, induced independently by the release of vasopressin systemically and in brain, may synergize to achieve coordinated physiological and behavioural response. A parallel example is to be found in the studies of cholecystokinin[12] and bombesin[13] both of which have been reported, after systemic injection, to reduce food intake in the fasted animal. Traditionally such effects are thought to involve satiety centres of the CNS, which modulate ingestive behaviour. However, Deutch[14] and his colleagues found that these peptides induce profound physiological changes in gut and suggest that the resulting malaise or strong visceral arousal is sufficient to account for the behavioural effects of these peptides. In support of this view, Deutch has reported that subcutaneously administered cholecystokinin or bombesin provides a sufficient visceral cue for conditioned taste aversion to develop.

Again, CNS activation, secondary to visceral arousal, may largely account for the behavioural effect of a peptide. The importance of the present article by Le Moal *et al.* is that it forces us to re-evaluate the peripheral effects of peptides in accounting for their effects on behaviour. In the past decade it has been exciting to discover that a large number of peptides, with established endocrine and physiological functions, exist in brain. Perhaps we have too readily assumed that their behavioural effects reflect direct actions on the CNS. ☐

100 years ago

Mr. Josceline Bagot and Mr. Drummond, of the Grenadier Guards, accompanied by Mr. T. Wright, the winner of the International Balloon Contest, went up in a balloon from the Crystal Palace on the 1st inst. at 1 p.m. When the ropes were loosed they ascended to the height of 5000 feet, and travelled slowly in a south-westerly direction for the distance of about eight miles. The balloon then suddenly sank, but ballast being thrown out, it rose again to 8000 feet, and traversed in the direction of Epsom. The aëronauts then descended in a field about a quarter of a mile from the Grand Stand, which they reached in time to witness the race for the Derby.

An opera performed on the stage of the Paris Grand Opera House was heard satisfactorily at the Rue Riche Opera House by a number of French officials a few days ago. The feat was performed with the new Ader telephone, of which the peculiarities have not been made public. The performance will be repeated at the International Exhibition.

from *Nature* **24**, 9 June, 132, 1881

Susan D. Iversen is Assistant Director of Research at the Department of Experimental Psychology, Cambridge, UK.

1 De Wied, D. & Bohus, B. *Nature* **212**, 1484 (1966).
2 De Wied, D. *Nature* **232**, 58 (1971).
3 De Wied, D. *Life Sci.* **19**, 685 (1976).
4 Swanson, L.W. *Brain Res.* **128**, 346 (1977).
5 Kovacs, G.L., Bohus, B. & Versteeg, D.H.G. *Neuroscience* **4**, 1529 (1979).
6 Rossor, M.N., Iversen, L.L., Hawthorn, J., Ang, V.T.Y. & Jenkins, J.S. *Lancet* ii, 1367 (1980).
7. Mason, S.T. & Iversen *Brain Res. Rev.* **1**, 107 (1979).
8. Bondareff, W., Mountjoy, C.Q. & Roth, M. *Lancet* i, 783 (1981).
9. McEntee, W.J. & Mair, R.G. *Science* **202**, 905 (1978).
10. Oliveros, J.C. *et al. Lancet* i, 42 (1978).
11. De Wied, D., Greven, H.M., Lande, S. & Witter, A. *Br. J. Pharmac.* **45**, 118 (1972).
12. Smith, G.P. & Gibbs, J. *Pharmac. Biochem. Behav.* **3**, (Suppl 1) 135 (1975).
13. Gibbs, J. *et al. Nature* **282**, 208 (1980).
14. Deutch, J.A. *Nature* **285**, 592 (1980).

0028-0836/81/240454-01$01.00

REVIEW ARTICLE

The traditional and a new version of the mouse *H–2* complex

Jan Klein, Antonio Juretić*, Constantin N. Baxevanis & Zoltan A. Nagy
Max-Planck-Institut für Biologie, Abteilung Immungenetik, Corrensstrasse 42, 7400 Tübingen, FRG

The H–2 *complex has traditionally been interpreted as a maze of regions, subregions and loci coding for different traits. The two main theses presented here are, first, that a single* H–2 *locus is pleiomorphic in that it controls several functions such as allograft rejection, cell-mediated lymphocytotoxicity, mixed lymphocyte reaction, immune response, immune suppression and restriction of T-cell specificity; and second, that the physiological function of the* H–2 *complex is to guide T lymphocytes in their function of distinguishing self from non-self, and that all other* H–2-*controlled traits are artificial derivatives of this basic function. These two theses lead to a new, simplified interpretation of the* H–2 *complex.*

ACCORDING to the traditional view, the major histocompatibility complex (MHC) is a chromosomal segment bounded by two loci originally chosen because their products are responsible for rapid rejection of foreign tissue grafts and are easily detected by serological methods[1,2]. Each locus in this segment is then considered part of the MHC, whether it is functionally related to the bounding loci or not. The organization of the *H–2* complex, the mouse MHC, is shown in Table 1. The complex is divided into four regions, *K; I, S* and *D*, and the *I* region is further divided into five subregions: *I-A, I-B, I-J, I-E* and *I-C* (the so-called *G* region, originally placed between *S* and *D*, has recently been demonstrated to be inseparable from the *S* region[3,4]). Each region or subregion controls several traits, and each trait is presumably controlled by a separate locus. The relationship between the various loci in each region or subregion is unknown but in no case have genetic recombinants been obtained by separating the individual loci. A brief description of the *H–2*-controlled traits is given below (see also Table 1).

H–2-controlled traits

Rejection of allografts. A piece of tissue or an organ taken from one mouse and transplanted to another mouse with a different *H–2* complex is destroyed by an immune process called the allograft reaction. The alloantigens which trigger this reaction are controlled by histocompatibility (*H*) loci which map in the *K, I-A, I-E* and *D* regions[5].

Cell-mediated lympholysis (CML). Thymus-derived (T) lymphocytes of one mouse can be induced by cells of another mouse with a different *H–2* complex to transform into cytolytic (killer) cells that can lyse appropriate target cells. This lysis is usually measured by the release of cytoplasmically bound isotopes such as ^{51}Cr. The loci controlling the CML-inducing alloantigens map in the same regions as the *H* loci and the two are usually considered to be identical.

Mixed lymphocyte reaction (MLR). Resting T lymphocytes of one mouse can be stimulated to proliferate by mixing them in tissue culture with lymphoid cells from another mouse with a different *H–2* complex. The proliferation can be measured by the incorporation of ^{3}H-thymidine into the DNA-synthesizing cells. The lymphocyte-activating determinants responsible for the MLR are controlled by *Lad* loci which map predominantly in the *I-A* and *I-E* subregions[6], although weak *Lad* loci have also been placed in the *K, I-J, I-C* and *D* regions[7–9].

Control of immune response. There are striking differences in the ability of different mouse strains to produce antibodies to a great variety of antigens[10]. Some strains are high responders which produce high levels of antibodies to a given antigen, others are low responders or nonresponders which produce low levels of or no antibodies to this antigen. The variation in the immune response ability is controlled by immune response (*Ir*) loci, many of which map in the *I* region of the *H–2* complex[11] (however, *Ir* loci mapping in the *K* region have also been reported[12,13]). In some cases, the nonresponder can be turned into a high responder by immunization with an antigen attached to another molecule (for example, methylated bovine serum albumin, MBSA[14]). In such situations, preimmunization with the antigen alone often suppresses the response to the subsequent challenge with the complexed antigen, a response that, without the pretreatment, is high[15]. The suppression is thought to be mediated by special immune suppression (*Is*) genes that map in the *I-A* and *I-C* subregions. Immune suppression is generally mediated by a subpopulation of T lymphocytes, the suppressor T cells (T_s), some of which carry a cell-surface alloantigen controlled by the *I-J* subregion[16]. The *I-J* antigens are detected by specific alloantibodies which abolish immune suppression when incorporated into some of the immune suppressive systems, presumably by blocking or killing the suppressor T cells.

Control of T–B collaboration. Although antibodies are produced by descendants of B lymphocytes, the differentiation of B cells into antibody-producing cells is triggered by the interaction of B lymphocytes with the helper T lymphocytes (T_h). This interaction is mediated by hypothetical 'interaction structures' controlled by *I*-region loci[17]. Originally, T cell–B cell compatibility of interaction structures was thought to be required for successful collaboration[18], but this conclusion was later challenged[19].

Restriction of T-cell specificity. Virus-infected cells stimulate the differentiation of precursor T cells into cytolytic T (T_c) lymphocytes capable of destroying infected target cells. The T-cell response to virus infection shows dual specificity: the stimulated cells are specific not only for the virus but also for H-2 molecules borne by the inducing cells and controlled by the *K* and *D* regions[20]. As a result the mature T_c lymphocytes lyse only target cells infected with the same virus and bearing the same K or D molecules as the stimulating cells. Similar H-2 restriction of T-cell activity has also been described for the T_h lymphocytes. To help B cells, the T_h cells must first be activated by an antigen presented to them by a macrophage-like (antigen-presenting) cell. The T_h cells also show dual specificity in that

---

* Permanent address: Department of Physiology, Faculty of Medicine, University of Zagreb, Zagreb, Yugoslavia.

0028–0836/81/240455—06$01.00

Table 1 The traditional version of the *H–2* complex

Region	*K*	*I*					*S*	*D*
Subregion		*I–A*	*I–B*	*I–J*	*I–E*	*I–C*		
Molecular product	K	A_α, A_β, A_e	?	?	E_α	?	SS, Slp	D, L
Trait (locus)								
Serologically detectable antigens (*H–2, Ia, SS, Slp*)	+ + + +	+ + + +	—	+	+ +	—	+ + + +	+ + + +
Rejection of allografts (*H*)	+ + + +	+ + +	—	—	+ +	?	—	+ + + +
Cell-mediated lympholysis (*H*)	+ + + +	+ + +	—	—	+ + +	?	—	+ + + +
Mixed lymphocyte reaction (*Lad*)	+ +	+ + + +	—	+	+ +	+	—	+ +
Control of immune response (*Ir*)	+	+ + + +	+ +	—	+ + +	—	—	—
(*Is*)	—	+ + + +	—	—	—	+ + + +	—	—
T_s-cell marker	—	—	—	+ + + +	—	—	—	—
Control of T–B cell collaboration	—	+ + + +	?	?	+ + +	?	—	—
Restriction of T_c-cell specificity	+ + + +	—	—	—	—	—	—	+ + + +
Restriction of T_h-cell specificity	—	+ + + +	?	?	+ + +	?	—	—
Control of complement activity	—	—	—	—	—	—	+ + + +	—

they recognize not only the antigen but also H–2-encoded molecules of the antigen-presenting cell. However, in contrast to the T_c cells, the molecules recognized by the T_h cells are controlled by loci in the *I* region[21].

Control of complement activity. Complement is a system of interacting serum proteins (mostly enzymes), activated in a coordinated fashion, which participate in immunological functions such as lysis of bacteria and macrophage chemotaxis. Complement-encoding loci have been mapped in the *S* region of the *H–2* complex[22].

Gene products

The usual way of detecting H–2-controlled molecules is to cross-immunize strains differing only in the *H–2* complex (*H–2* congeneic lines) or only in part of it (*H–2* recombinant strains), and to produce antibodies against these molecules. These antibodies can then be analysed by any of the known serological methods and individual H–2 determinants can thus be defined. Serologically detectable molecules controlled by all *H–2* regions or subregions except *I–B* and *I–C* have now been identified. The determinants borne by the K and D molecules are referred to as H–2 antigens, those controlled by the *I* subregion are termed *I* region-associated, or Ia antigens.

The *K* region codes for at least one polypeptide chain with a molecular weight (MW) of 45,000 (ref. 23), which is non-covalently associated with the small polypeptide β_2-microglobulin encoded by a locus on chromosome 2 (ref. 24) (the *H–2* complex is located on chromosome 17). The *D* region controls at least two polypeptide chains, D and L, each of MW 45,000 and each associated with β_2-microglobulin polypeptide chains.

The *I* region codes for a minimum of four polypeptide chains A_α, A_β, A_e (E_β) and E_α; the α chains have a molecular weight of ~32,000, the β chains of ~28,000[25]. The chains are synthesized separately in the cytoplasm but combine to form heterodimers before they integrate into the plasma membrane: the A_α combines with A_β to form an A molecule and the A_e (so designated because the encoding locus maps in the *I–A* subregion) combines with E_α to form an E molecule. Some inbred strains carry a mutation affecting the E_α-encoding locus which results in the absence of E_α chains in the cytoplasm; strains carrying this mutation do not express the E molecule on the cell surface, as in the absence of the α chain, the β chains cannot integrate into the membrane[26]. The products of the *I–B*, *I–J* and *I–C* subregions have not been identified.

The *S* region codes for two molecules, Ss and Slp, both of which correspond structurally to the C4 component of the complement cascade but only one of which (Ss) has been demonstrated to have C4 activity in the functional test[27].

This traditional interpretation of the *H–2* complex grew out of total ignorance of the physiological function of the *H–2* loci. It is understandable that in a situation where seemingly unrelated traits were discovered to be *H–2* controlled, the complex seemed to immunologists to be a maze of separate loci, regions and subregions, the interrelationships of which remained unclear. However, now that we have some idea about the *H–2* function (see below), we should ask whether the complex has not been over-interpreted and whether it cannot be viewed in a simpler way. Here we propose a new interpretation of the *H–2* complex which is simpler and unifies the loci. For some, this article may merely be a formulation of what they have thought for some time; for others it may represent a radical departure from their own concepts; for us, it is the only interpretation that makes sense.

Unification of *H–2* loci

One of the main theses of this article is that a single *H–2* locus controls a variety of traits—in other words, that the loci are pleiomorphic. There is much data supporting this idea, of which we will cite here only a few selected examples.

Unity of loci in the *K* and D regions. These two regions contain loci which govern rapid allograft rejection, strong CML, weak MLR, immune response to a few antigens and specificity of cytolytic T lymphocytes[1,2]. We propose that all these functions are carried out by the serologically detectable, 45,000-MW molecules controlled by the *H-2K* and *H-2D* (*H-2L*) loci in these two regions. The main evidence for this conclusion comes from the study of *H–2* mutations; over 24 mutations mapping in the *K* (most of them) or the *D* (a few) region have been described[28]. All have been demonstrated to affect the *H-2K*, *H-2D or H-2L* and no other loci in the genome. Several are known to have resulted in amino acid substitutions in the serologically detectable, 45,000-MW polypeptide chain[29]. Functional tests have revealed that a single mutation in the *H-2K*, *H-2D* and *H-2L* locus can generate new antigenic determinants which are detectable by serological methods, lead to graft rejection, stimulate CML and MLR, change the immune response pattern and affect the specificity of cytolytic T lymphocytes directed against virus-infected cells[28]. There is, therefore, every reason to believe that all these traits are the function of a single molecule—the 45,000-MW, serologically detectable molecule—and there is thus no need to postulate separate loci for the individual traits.

Unity of loci in the *I* region. We shall deal first with the *I–A* and *I–E* subregions—the two subregions known to control serologically detectable Ia antigens. All the *I–A*- or *I–E*-associated functions are carried out by the A and the E molecules and we shall now discuss the evidence for the identity of the individual loci.

Identity of *H* and *Ia* loci. A unique opportunity to test the relationship of the *H* loci governing graft rejection *in vivo* or CML *in vitro*, and the *Ia* loci has been provided by the discovery that some mouse strains carry a nonexpressor allele at the E_α locus which prevents the appearance of E molecules on the cell surface[26]. We have exploited the existence of nonexpressor E_α alleles (which we designate E_α^0; the expressor alleles are designated E_α^+) in the following way[30,31]. We selected from the collection of available mouse strains four pairs in which the members of each pair were genetically identical except for a short segment of the *I* region, including the *I–E* subregion. Furthermore, the strain combinations were such that in each pair one member carried the E^0_α and the other member the E^+_α allele. When we then attempted to generate primary CML with cells derived from these strains, we succeeded in the following combinations: B10.A(4R) anti-B10.A(2R), B10.D2(R107) anti-B10.A(3R), B10.S(7R) anti-B10.S(9R) and D2.GD anti-B10.D2(R101). All attempts to generate CML in the reciprocal direction (such as B10.A(2R) anti-B10.A(4R), B10.A(3R) anti-B10.D2(R107)) failed. Examination of the genetic constitution of the eight strains used (Table 2) reveals that positive CML reaction occurred in combinations in which the responder strain carried the E^0_α allele (and hence did not express the E molecule) and the stimulator the E^+_α allele (and hence expressed the E molecule); no CML could be generated against cells that lacked the serologically detectable I–E molecule. This correlation between generation of CML and the presence of the E molecule on stimulator cells can hardly be a coincidence; most likely the serologically detectable I–E molecule is the one that generates the CML and is thus the product of the *I*-region *H* gene. A similar case can also be made for the *H* locus mapping in the *I–A* subregion[32]. Here the CML reaction directed against a product of this subregion can be blocked by Ia-specific antibodies[33], and a mutation affecting the serologically detectable Ia antigens also affects the *H* locus[34]. The combined data thus best fit the interpretation that the *H* and *Ia* loci in the *I–A* and *I–E* subregions are identical.

Identity of *Lad* and *Ia* loci. The evidence for equating the *Lad* loci, which govern the MLR, with the *Ia* loci, responsible for the serologically detectable determinants, is of a similar nature to that for the *H* and *Ia* loci: the reaction can be inhibited by anti-Ia sera[35] and the same mutant mentioned above also affects MLR[34]. In addition, the Lad and Ia determinants are known to have the same tissue distribution, being particularly abundant on B lymphocytes and macrophages, or macrophage-like cells, and scarce on most other cells[36].

There are two additional, rather peripheral pieces of evidence which nevertheless argue strongly in favour of our thesis. The first concerns an older observation[37] that in strain combination B10.A(2R) and B10.A(4R), MLR occurs only in one direction (B10.A(4R) anti-B10.A(2R)). This unidirectionality, which has long puzzled immunologists, can now be explained simply by the fact that the two strains differ only in the *I*-region segment around the *I–E* subregion and that the B10.A(2R) strain expresses the E molecule whereas the B10.A(4R) strain does not. The occurrence of MLR thus again correlates with the expression of serologically detectable E molecules on the surface of the stimulating cells.

The second piece of evidence has been provided by Fathmann *et al.*[38] who observed that T lymphocytes of strain A (haplotype $H\text{–}2^a$ with alleles $K^k, I\text{–}A^k, I\text{–}B^k, I\text{–}J^k, I\text{–}E^k, I\text{–}C^d, S^d$ and D^d), when sensitized to (A × C57BL/6)F$_1$ cells respond on restimulation more strongly to the F_1 than to the C57BL/6 cells (haplotype $H\text{–}2^b$ with alleles K^b, $I\text{–}A^b$, $I\text{–}B^b$, $I\text{–}J^b$, $I\text{–}E^b$, $I\text{–}C^b$, S^b and D^b). The authors explain this result by postulating the existence of "hybrid determinants" in the *H–2* heterozygous cells, but a simpler explanation is possible. In the A anti-(A × C57BL/6)F$_1$ combination, two kinds of cell are activated: those which recognize A^b_β determinants (in combination with A^b_α or A^k_α) and those recognizing E^b_β determinants (in combinations with E^k_α). When tested against the C57BL/6 parent which carries the null allele at E_α and hence does not express cell-surface E^b_β, only the anti-A^b_β cells (not the anti-E^b_β cells) can react. For this reason the response against the F$_1$ hybrid is stronger than that against C57BL/6 cells. This observation again strongly argues that the lymphocyte-activating determinants are carried by the same molecules that bear the serologically detectable Ia antigenic determinants.

Table 2 *H–2* genotypes of recombinant strains used in the study of phenotype–genotype relationships

Strain	*H–2* haplotype	Alleles at *H–2* loci K	A_β	A_α	E_β	E_α	D
B10.D2(R101)	*g1*	*d*	*d*	*d*	*d*	*b*	*b*
B10.GD	*g2*	d	d	d	(*d*/*b*)	(*b*)	*b*
B10.A(2R)	*h2*	*k*	*k*	*k*	*k*	*k*	*b*
B10.A(4R)	*h4*	*k*	*k*	*k*	(*k*)	(*b*)	*b*
B10.A(3R)	*i3*	*b*	*b*	*b*	*b*	*k*	*d*
B10.D2(R107)	*i7*	*b*	*b*	*b*	(*b*)	(*b*)	*d*
B10.S(7R)	*t2*	*s*	*s*	*s*	(*s*)	(*s*)	*d*
B10.S(9R)	*t4*	*s*	*s*	*s*	*s*	*k*	*d*

The strains are listed in pairs where members of each pair differ only at the E-encoding loci. Alleles in parentheses are not expressed on the cell surface. The loci are listed according to the simplified interpretation of the *H–2* complex given in the text.

Identity of *Ir* (*Is*) and *Ia* loci. While the only way to study *Ir* genes was to immunize mice and test their sera for the presence of specific antibodies, the question of the *Ir*-gene product was almost unanswerable. Fortunately, however, an assay is now available in which part or the entire immune response process occurs in tissue culture and is thus amenable to cellular and molecular manipulation. This 'T-cell proliferation assay'[39] consists of immunizing mice with a given antigen, collecting immune T cells several days after the immunization, and culturing these together with macrophages and the same antigen as that used for immunization. The macrophages presumably present the antigen to the T cells which then respond by proliferation and this can be measured by the incorporation of ^{3}H-thymidine. Numerous tests done using this assay have invariably proved that the degree of T-cell proliferation *in vitro* correlates well with the level of antibody production to the particular antigen *in vivo*: high responders defined by the antibody assay are also high responders *in vitro* and T cells from strains that produce low levels of or no antibodies *in vivo* proliferate poorly or not at all *in vitro*[40]. This correlation is not surprising in view of the fact that helper T cells are necessary to initiate B-lymphocyte differentiation, leading to antibody-producing cells.

We used the T-cell proliferation assay to test the relationship between the *Ir* and *Ia* genes in the following way[41]. Three antigens were selected, the response to which is known to be under *Ir* control: synthetic polypeptide of randomly assembled amino acids glutamic acid and alanine (GA), another synthetic polypeptide of the randomly assembled amino acids glutamic acid, lysine and tyrosine (GLT), and a natural protein lactate dehydrogenase B (LDH$_B$). The antibody response to GA is known to be controlled by an *Ir* gene that maps in the *I–A* subregion[42]. We have tested the GA response of several strains in the T-cell proliferation assay and found complete agreement with the *in vivo* data: in this assay the response is also controlled by the *I–A* subregion. We then repeated the experiments but added to the culture monoclonal antibodies specific for Ia-bearing molecules. We found that the proliferation of T cells derived from high-responder strains was inhibited by the Ia-specific antibodies but only by those reacting with the A molecule of the strain from which the cells in the culture were derived; antibodies specific for the E molecule of the same strain or those specific for the H-2K and H-2D molecules had no effect on the reaction. This experiment and similar ones done by others[43–45] provides evidence that the Ia-bearing molecules are involved in an important way in the control of the immune response to antigens. Even more revealing are the data obtained with the second antigen. The antibody response to GLT has been reported to be controlled by two *Ir* genes[46], one mapping in the *I–A* and the other in the *I–E* subregion. In addition, the response to this antigen shows a phenomenon known as *Ir*-gene complementation[47]. It has been observed that some of the recombinants derived by crossing of two nonresponder strains are high responders. This has been explained by assuming that two genes are involved in the response and that neither of these genes alone can effect high responsiveness but that the combined action (complementation) of the two genes does result in high responsiveness.

We tested the response to GLT in the T-cell proliferation assay and again confirmed the mapping data obtained by *in vivo* antibody testing: there are indeed two genes involved, mapping respectively in the *I-A* and *I-E* subregions[41]. However, when we tried to inhibit the proliferation with Ia-specific monoclonal antibodies, we discovered that such inhibition was possible only with the E molecule-specific and not with the A molecule-specific antibodies. We explain these data as follows. In contrast to GA, the immune response to GLT is controlled by the E molecule, more specifically by the gene coding for the E_β chain, which is located in the *I-A* subregion. However, because the control is dependent on the expression of the E_β chain on the cell surface and as this depends on the presence of the E_α^+ gene, which maps in the *I-E* subregion, at the phenotype level the response to GLT seems to be controlled by two genes, one in the *I-A*, the other in the *I-E* subregion. In some recombinants derived from two nonresponder strains, a high responder E_β allele of one parent, which was not phenotypically expressed in the parent because it was combined with an E_α^0 allele, combines with an E_α^+ gene, becomes expressed, and results in responsiveness.

The combined data on GA and GLT responsiveness indicate that *Ir*-gene effects can be fully explained by assuming that the control of immune responsiveness occurs through the serologically detectable A or E molecules, and thus that the Ia-bearing molecules are the long-sought products of the *Ir* genes. The response to some antigens (for example, GA) is controlled by the A molecule, whereas that to other antigens (for example, GLT) is controlled by the E molecule. This conclusion is supported by the study of responsiveness to LDH_B, which also reveals additional complexity and which we shall discuss later.

We propose that the *Is* genes represent situations in which two H-2 molecules become involved in the response: one controls the generation of helper T cells and the other of suppressor T cells. The latter then act on the former and suppress the response. A precedent for this type of control exists in the response to LDH_B (see below). The proposition can be tested experimentally using monoclonal Ia- or H-2-specific antibodies.

Identity of loci controlling T_h-cell specificity with *Ia* loci. This genetic relationship follows from our experiments and others[43–45] in which the T_h-cell function was blocked by Ia-specific antibodies. The attachment of antibodies to the Ia-bearing molecules on antigen-presenting cells presumably prevents T_h cells from recognizing these molecules and therefore from being stimulated by the antigen. Furthermore, as the genetic restrictions of T–B cell collaboration *in vivo* correlate perfectly with those determined by the T-cell proliferation assay *in vitro*, the postulated interaction structures[17] are nothing but the Ia-bearing molecules.

In summary, all the different functions attributed to the *I-A* and *I-E* subregions of the *I* region have a common denominator in the serologically detectable, Ia determinant-bearing molecules. It is these molecules that are the products of the loci controlling these functions and hence the loci encoding the different traits are all identical.

Problematic subregions of the *I* region

We shall discuss the remaining three *I* subregions separately as they each present a different problem in their interpretation.

I-B subregion. This chromosomal segment was originally defined in the manner shown in Fig. 1. Here, the *H-2*^{a}/*H-2*^{b} heterozygote gave rise to two reciprocal haplotypes, *H-2*^{i5} and *H-2*^{h4}. The *H-2*^{a} haplotype confers low responsiveness to antigens such as LDH_B or mouse IgG2a myeloma protein; the *H-2*^{b} and both recombinant *H-2* haplotypes confer high responsiveness to these antigens[48,49]. This observation was interpreted as indicating the existence of a new *I-B* subregion separated from the *I-A* subregion by the recombinational event leading to the *H-2*^{h4} haplotype and from the *I-J* subregion by crossing-over resulting in the *H-2*^{i5} haplotype (see Fig. 1). The control of immune responsiveness to the LDH_B and IgG2a, as well as a few other antigens, has remained the only known function of the *I-B* subregion: this subregion does not affect MLR or CML, does not influence graft rejection, and does not control any serologically detectable product.

We have tested again the responsiveness to LDH_B and to the IgG2a myeloma protein using the *in vitro* T-cell proliferation assay and found that the *H-2*^{a} is a nonresponder haplotype and that the *H-2*^{b}, *H-2*^{h4} and *H-2*^{i5} are high responder haplotypes just as in the antibody assay[50]. However, by using the antibody-inhibition method we found that the responsiveness of the *H-2*^{b} haplotype could be completely abolished by the addition to the culture of monoclonal antibodies specific for the A molecule. This observation suggests that the generation of T_h cells in response to LDH_B or IgG2a is controlled by the *I-A* rather than by the *I-B* subregion. This discrepancy between formal genetic analysis and functional testing was eliminated when we performed similar experiments with the H-2^a cells. The addition of E molecule-specific antibodies to the culture of H-2^a cells, which are normally nonresponders, transformed them into high responders. Instead of blocking, the antibodies in this case enhanced the response. As we obtained a similar conversion of nonresponders into high responders by the removal from the culture of suppressor T cells, we explained the data as follows. In mice expressing the A and the E molecules, the response to LDH_B and IgG2a is controlled by both these products but the E molecule controls the generation of suppressor T cells which then inhibit helper T cells, the generation of which is controlled by the A molecule. The coating of the E molecule with specific antibodies prevents the recognition of these molecules and thus the induction of suppression; the proliferation of T_h cells is then unopposed. In strains that do not express cell-surface E molecules (for example, the H-2^b strains) only T_h cells are generated and the mice then show a high responder phenotype. This interpretation completely disposes of the *I-B* subregion as the response to LDH_B and similar antigens is fully explained by the interaction of cells recognizing I-A- and I-E-controlled molecules. We suggest that the *I-B* subregion is a genetic artefact and as such should be removed from the maps of the *H-2* complex. At the same time, the data strongly support the concept that the products of the *Ir* genes are the Ia-bearing molecules.

***I-J* subregion.** Although it is often stated that the *I-J* subregion controls immune suppression[51], we are not aware of any convincing data which support this contention. The experiments with I-J-specific antibodies have only demonstrated that some suppressor T cells express I-J-controlled determinants on their surfaces and that I-J determinants are present on Ts-derived factors. However, this finding could mean that I-J determinants are merely markers of suppressor T cells, just as the so-called

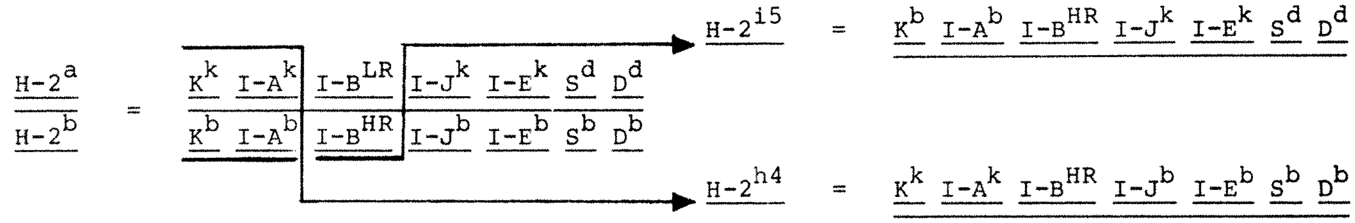

H-2^i5 = K^b I-A^b I-B^HR I-J^k I-E^k S^d D^d
H-2^a / H-2^b = K^k I-A^k I-B^LR I-J^k I-E^k S^d D^d / K^b I-A^b I-B^HR I-J^b I-E^b S^b D^b
H-2^h4 = K^k I-A^k I-B^HR I-J^b I-E^b S^b D^b

Fig. 1 Definition of the *I-B* subregion by formal genetics. The figure shows the derivation of two *H-2* recombinants, *H-2*^{i5} and *H-2*^{h4}, from an *H-2*^{a}/*H-2*^{b} heterozygote. To explain the pattern of responsiveness to LDH_B or mouse IgG2a myeloma protein (indicated by the HR and LR symbols which represent 'high' and 'low' responsiveness, respectively), an *I-B* subregion is postulated. The directions of the genetic exchanges leading to the recombinants are indicated by arrows.

Lyt antigens are markers of certain other T-cell subpopulations[52]. It may be, therefore, that the J molecule is unrelated to the A and E molecules; in the absence of hard data regarding the I-J function it is better to remain uncommitted. ***I–C* subregion.** This portion of the *I* region was once thought to control serologically detectable molecules[53] but this finding could not be substantiated[54]. There are, however, data in the literature which suggest that the *I–C* subregion controls the generation of suppressor T cells in MLC[55]. Also, the *Is* genes may require the existence of a separate *I–C* subregion[47]. However, the evidence for the *I–C* subregion is so weak that at present it is better to leave the subregion out of the *H–2* maps until its existence is fully established.

S region

This chromosomal segment is strikingly different from the rest of the *H–2* complex: it does not control any of the traits associated with the *K*, *D* or *I* regions, and complement activity, the only trait that it does control, is not affected by the other regions. This gives the impression that the *S* region is an alien immigrant into the *H–2* complex and has functionally nothing to do with the complex itself. (The argument that the *S* region is part of the *H–2* complex because, like the *H–2* proper, it affects immunological functions is absurd because then all loci affecting immunological functions would have to be considered as members of this complex.)

Function of the *H–2* loci

We have collected evidence which indicates that each region or subregion of *H–2* (with the exception of *D*) contains only one locus which controls all the traits attributed to this chromosomal segment. The question which then arises is which of the traits reflect the true function of *H–2*. The answer to this question is not as difficult as it might seem; several of the traits listed in Table 1 are artefacts that cannot possibly have any physiological role. Mice do not normally encounter tissue grafts or lymphocytes from other mice and thus allograft reaction, CML and MLR cannot have any function in an intact individual. On the other hand, immune responsiveness, immune suppression, T–B cell interaction, and the restriction of T-cell specificity are physiologically important phenomena which must be related to the natural *H–2* function. Because all these phenomena have a common denominator in that they reflect the recognition by T cells of an antigen in the context of H-2-controlled molecules, the role of *H–2* must be to provide this context for antigen recognition. This dual recognition may be the only way for organisms to distinguish non-self from self.

The basis for the *Ir*-gene phenomenon is the fact that certain combinations of self (H–2 molecules) and non-self (antigen) fail to stimulate T lymphocytes. This failure could have at least two reasons: in some situations the particular antigen/*H–2* combination is not recognized by T cells for reasons which we have no space to discuss here; in other situations helper T cells are stimulated but, at the same time, so are suppressor T cells which then prevent the T_h cells from performing their function.

The requirement for simultaneous recognition of antigen and H–2 molecules restricts the specificity of the T lymphocyte—this restriction is probably the real function of the *H–2* complex. Molecules known to perform this function are all heterodimers consisting of one highly variable (K, D, L, A_β and E_β) and one relatively constant (β_2-microglobulin, A_α and E_α) polypeptide chain. The variability of K, D, L and A_β chains is generally recognized; that of the E_β chain deserves a special comment. Serological studies have allowed the definition of only three E_β alleles[56], thus it is often thought that the E_β-encoding locus is less polymorphic than the A_β locus, at which different alleles are known to be present in every independent haplotype tested. However, using a sensitive CML assay, we have recently obtained data which indicate that E_β chains encoded by the different independent haplotypes differ antigenically from one another. This observation suggests that the E_β chains are as variable as the A_β or the K and D chains[30,31]. The failure to distinguish the E_β chains by serological methods can probably be attributed to the fact that many independent haplotypes carry the E_α^0 allele and that the strains bearing this haplotype do not express the E molecule on the cell surface. One should therefore search for haplotype-specific E_β determinants in cells that allow the expression of the E molecule—that is, in appropriate recombinants or *H–2* heterozygotes.

The question of how the H–2 molecules actually carry out their function remains largely open. The two principal possibilities are, first, that the H–2 molecules physically interact with the antigen on the antigen-presenting cells and that this interaction then gives the impression of dual specificity of the T cells. Second, that the basis for the dual specificity is the expression by each T cell of two kinds of receptor, one for the antigen and the other for the H–2 molecules[57]. The former hypothesis presupposes a receptor function for the H–2 molecules, a supposition which is not only contradicted by the available data but is also difficult to justify on theoretical grounds. These considerations leave us with the latter hypothesis and hence with the assumption that there are separate T-cell receptors for antigens and H–2; the two receptors, however, need not be distinct.

A simplified view of the *H–2* complex

If we accept the proposition that the function of the *H–2* complex is to guide T lymphocytes in distinguishing between self and non-self and that the other traits currently attributed to the H–2 molecules are a reflection of this basic function, we arrive at the following model of *H–2*.

Locus	*K*	A_β	A_α	E_β	E_α	*D*	*L*
Molecule	45K	28K	32K	28K	32K	45K	45K
Functional class	I	II	II	II	II	I	I

(The order of the loci enclosed in the bracket is unknown. K represents 1,000 MW.) Here the complex is divided into two classes of loci, class II composed of A_α-, A_β-, E_α- and E_β-encoding loci (suitable designations of these loci have not yet been agreed upon). The class I loci differ from those of class II in that the former guide cytolytic and the latter regulatory (helper and suppressor) T lymphocytes, although both classes operate on the same principle.

In this functional model of the *H–2* complex, we dispose of regions and subregions and retain only the term 'locus'. The division of the complex into regions and subregions is a historical artefact that confuses the picture, which in reality is probably rather simple. We eliminate from the complex the *B* locus for which we do not find any evidence, and postpone the inclusion of the *J* and *C* loci until convincing data are obtained for their functional relatedness to the complex (*J*) or for their existence (*C*). We also exclude from this definition of the *H–2* complex the complement-encoding loci which are not functionally related to the class I or class II loci. Although it is useful to list the *J* and the complement loci as markers on *H–2* maps, these loci should not be considered as part of the *H–2* complex proper. On the other hand, evidence is accumulating which indicates that the *Qa* and *Tla* loci, presently outside the boundaries of the *H–2* complex[58], are in fact part of the complex. Structurally and perhaps also functionally they resemble the class I loci. It is also possible that additional class I loci exist in the vicinity of the *K* locus and additional class II loci in the *A*-locus cluster. However, inclusion of these loci in the *H–2* map should be postponed until they are clearly established as distinct from all existing loci.

We do not claim that this simplified scheme is the final version of the *H–2* complex but we do believe that our interpretation explains the available data much better than the traditional model. We suggest that in H–2 studies the time has come to change the paradigm.

We thank Ms Rosemary Franklin for help in preparing this manuscript.

1. Klein, J. *Biology of the Mouse Histocompatibility-2 Complex* (Springer-Verlag, New York, 1975).
2. Klein, J. *Science* **203,** 516–521 (1979).
3. Huang, C.-M. & Klein, J. *Immunogenetics* **11,** 605–616 (1980).
4. Ferreira, A., David, C. S. & Nussenzweig, V. *J. exp. Med.* **151,** 1424–1435 (1980).
5. Klein, J. *Springer Semin. Immunopath.* **1,** 31–49 (1978).
6. Nabholz, M. & Miggiano, V. C. in *B and T Cells in Immune Recognition* (eds Loor, F. & Roelants, G. R.) 261–289 (Wiley, Chichester, 1977).
7. Klein, J. *et al. Cell. Immun.* **4,** 442–446 (1972).
8. Okuda, K., David, C. S. & Shreffler, D. C. *J. exp. Med.* **146,** 1561–1573 (1977).
9. Lonai, P. & McDevitt, H. O. *Immunogenetics* **4,** 33–35 (1977).
10. Benacerraf, B. & McDevitt, H. O. *Science* **175,** 273–279 (1972).
11. McDevitt, H. O. *et al. J. exp. Med.* **135,** 1259–1278 (1972).
12. Zaleski, M. & Klein, J. *J. exp. Med.* **145,** 1602–1606 (1977).
13. Maron, R. & Cohen, I. R. *Nature* **279,** 715–716 (1979).
14. McDevitt, H. O. *J. Immun.* **100,** 485–492 (1968).
15. Kapp, J. A. *et al. J. exp. Med.* **140,** 648–659 (1974).
16. Murphy, D. B. *et al. J. exp. Med.* **144,** 699–712 (1976).
17. Katz, D. H. *et al. Proc. natn. Acad. Sci. U.S.A.* **70,** 2624–2628 (1973).
18. Katz, D. H., Hamaoka, T. & Benacerraf, B. *J. exp. Med.* **137,** 1405–1408 (1973).
19. Heber-Katz, E. & Wilson, D. B. *J. exp. Med.* **142,** 928–935 (1975).
20. Zinkernagel, R. M. & Doherty, P. C. *Nature* **248,** 701–702 (1974).
21. Sprent, J. *J. exp. Med.* **147,** 1159–1174 (1978).
22. Shreffler, D. C. *Immun. Rev.* **32,** 140–167 (1976).
23. Nathenson, S. G. *et al.* in *Birth Defects*, Vol. 14, No. 2 (eds Lerner, R. A. & Bergsma, D.) 566–579 (Liss, New York, 1978).
24. Michaelson, J. *Immunogenetics* **13,** 167–171 (1981).
25. Uhr, J. W. *et al. Science* **206,** 292–297 (1979).
26. Jones, P. P., Murphy, D. B. & McDevitt, H. O. *J. exp. Med.* **148,** 925–939 (1978).
27. Ferreira, A., Nussenzweig, V. & Oigli, I. *J. exp. Med.* **148,** 1186–1197 (1978).
28. Klein, J. *Adv. Immun.* **26,** 56–146 (1978).
29. Nairn, R., Yamaga, K. & Nathenson, S. G. *A. Rev. Genet.* **14,** 241–277 (1980).
30. Juretić, A., Nagy, Z. A. & Klein, J. *Nature* **289,** 308–310 (1981).
31. Juretić, A., Nagy, Z. A. & Klein, J. *Immunogenetics* (submitted).
32. Klein, J. *et al. J. exp. Med.* **143,** 1439–1452 (1976).
33. Klein, J., Chiang, C. L. & Wakeland, E. K. *Immunogenetics* **5,** 445–451 (1977).
34. McKenzie, I. F. C. *et al. J. exp. Med.* **150,** 1323–1328 (1979).
35. Meo, T. *et al. Transplantn Proc.* **7,** 127–129 (1975).
36. Hämmerling, G. *Transplantn Rev.* **30,** 64–82 (1976).
37. Lozner, E. C. *et al. Science* **183,** 757–759 (1974).
38. Fathman, C. G., Watanabe, T. & Augustin, A. *J. Immun.* **121,** 259–264 (1978).
39. Alkan, S. S. *Eur. J. Immun.* **8,** 112–118 (1978).
40. Schwartz, R. H. & Paul, W. E. *J. exp. Med.* **143,** 529–540 (1976).
41. Baxevanis, C. N. *et al. Immunogenetics* **11,** 617–628 (1980).
42. Merryman, C. F. & Maurer, P. H. *J. Immun.* **116,** 739–742 (1976).
43. Shevach, E. M., Paul, W. E. & Green, I. *J. exp. Med.* **136,** 1207–1221 (1972).
44. Schwartz, R. H. *et al. J. Immun.* **117,** 531–540 (1976).
45. Schwartz, R. H. *et al. Proc. natn. Acad. Sci. U.S.A.* **75,** 2387–2391 (1978).
46. Merryman, C. F. & Maurer, P. H. *Immunogenetics* **1,** 549–559 (1975).
47. Dorf, M. E. *Springer Semin. Immunopath.* **1,** 171–206 (1978).
48. Lieberman, R. *et al. J. exp. Med.* **136,** 1231–1240 (1972).
49. Melchers, I., Rajewsky, K. & Shreffler, D. C. *Eur. J. Immun.* **3,** 754–761 (1973).
50. Baxevanis, C. N., Nagy, Z. A. & Klein, J. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
51. Murphy, D. B. *Springer Semin. Immunopath.* **1,** 111–131 (1978).
52. Cantor, H. & Boyse, E. *J. exp. Med.* **141,** 1376–1389 (1975).
53. Shreffler, D. C. & David, C. S. *Adv. Immun.* **20,** 125–195 (1975).
54. Cullen, E. S. *et al. Immunogenetics* **11,** 535–547 (1980).
55. Rich, S. S., David, C. S. & Rich, R. T. *J. exp. Med.* **149,** 114–126 (1979).
56. Klein, J. *et al. Immunogenetics* **6,** 489–512 (1978).
57. Zinkernagel, R. M. & Doherty, P. C. *Nature* **251,** 547–548 (1974).
58. Flaherty, L. in *Origins of Inbred Strains* (ed. Morse, H. C.) 409–422 (Academic, New York, 1978).

ARTICLES

Lead isotopic composition of Hercynian granitic K-feldspars constrains continental genesis

Annie Michard Vitrac*, Francis Albarede* & Claude J. Allègre

Laboratoire de Géochimie-Cosmochimie, Institut de Physique du Globe et Département des Sciences de la Terre,
4 Place Jussieu, 75230 Paris 5, France

Lead isotope compositions have been determined in many K-feldspars from granitic rocks emplaced during the Variscan (Hercynian) orogeny sensu lato *(around 300 Myr) in various parts of Europe and the eastern US. The ratios show a regular geographical pattern with* μ *contours arched around southern France and parallel to the Atlantic coast of the US. This distribution parallels predrift palaeogeographical zones during the Palaeozoic. The lead isotope pattern of Variscan granitic rocks is similar to that of some orogenic volcanic suites but is better explained by crustal melting than by an Hercynian mantle–crust material mixing. The continental crust from these areas must be older than 3,000 Myr. Large scale fluid circulation in the crust during orogenic processes apparently explains the observed geographical patterns of lead isotopes.*

WITHIN the general problem of formation and evolution of the continental crust, the nature of the Hercynian (or Variscan) European orogeny has been viewed as resulting from pure ensialic processes[1,2], as a former active continental margin[3] or as a product of continental collision[4,5]. An associated parameter is the degree of upper crust–lower crust–mantle material involvement in orogenic magmatism as it gives a direct insight on the nature of source areas affected by significant heat anomalies.

Since the early work of Hurley *et al.*[6] and Patterson and Tatsumoto[7], radiogenic isotopes have been shown to be powerful geological tracers for the study of the origin of continental crust. Studies over various geological provinces[8–11] showed that such an isotopic tracing technique strongly constrains the models proposed for orogenic processes. Here we consider the determination of lead isotopes in K-feldspars of ~300 Myr granitoids from western Europe and the Eastern US.

Because K-feldspars have fairly low U/Pb ratios[7], their lead isotopic composition is widely assumed to approximate that of the magma from which the crystals precipitated. In addition to allowing comparison with other studies of initial lead the analytical results were expected to provide new information about the geographical patterns shown by lead isotopes in the Hercynian orogeny and about the behaviour of the U/Pb system in the crust during orogenic processes. (Analytical techniques, experimental results and significant parameters are presented in Table 1.)

Lead isotopic provinces

The isotope ratios of ^{206}Pb/^{204}Pb, ^{207}Pb/^{204}Pb and ^{208}Pb/^{204}Pb vary between 18.04–18.65, 15.55–15.74 and 38.0–39.1 (Table 1, Fig. 1). As pointed out previously[8,12,13], the lead isotopic composition of feldspars is virtually independent of rock type in any area except for some rare ^{206}Pb/^{204}Pb variations discussed below.

The three lead isotopic ratios show a clear geographical distribution (Fig. 2), with the most radiogenic samples concentrated in southern France and the least radiogenic in southern Germany, northern France, Cornwall, Morocco and to the west of the Iberian Peninsula. Each of these regions usually gives fairly consistent results at a smaller scale than was observed by Zartman[13] in the western US.

Anomalously high ^{206}Pb/^{204}Pb ratios and young model ages may be explained by anomalously high U/Pb ratios in the corresponding feldspar samples. The correction for *in situ* U decay from the measured U and Pb concentrations is illusive as U is highly mobile in subsurface conditions[14,15], and U/Pb fractionation during sample leaching is an additional source of uncertainty. Within each region, anomalous ^{206}Pb/^{204}Pb ratios

---

*Present address: Centre de Recherches Petrographiques et Géochimiques and Ecole Nationale Superieure de Géologie, 54500 Vandoeuvre, France.

0028–0836/81/240460—05$01.00

Table 1 $^{206}Pb/^{204}Pb$, $^{207}Pb/^{204}Pb$, $^{208}Pb/^{204}Pb$ ratios for K-feldspars and plagioclases (Pl) from Hercynian granitoids classified by region and in granulitic–xenolith whole rocks from Massif Central

Sample	$^{206}Pb/^{204}Pb$	$^{207}Pb/^{204}Pb$	$^{208}Pb/^{204}Pb$	Single stage model μ	Single stage model K	Single stage model age (Myr)
Pyrénées						
Ml 20	18.400	15.681	38.25	9.645	3.943	282
Ml 13	18.373	15.713	38.59	9.713	4.120	339
	18.386	15.705	38.59			
	18.365	15.695	38.55			
Ml 12	18.380	15.694	38.56	9.673	4.094	312
Ml 22	18.580	15.708	38.57	9.679	3.987	187
Ml 34	18.447	15.704	38.58	9.687	4.069	278
Ml 57	18.420	15.692	38.54	9.665	4.062	281
Ml 32	18.453	15.717	38.66	9.711	4.101	288
	18.469	15.719	38.60			
Ml 66 Pl	18.472	15.695	38.51	9.665	4.019	248
	18.459	15.702	38.54			
Ml 64 Pl	18.436	15.710	38.60	9.699	4.083	292
Qt 04 Pl	18.398	15.673	38.46	9.630	4.034	274
Qt 04	18.428	15.683	38.48	9.646	4.028	263
Qt 08	18.455	15.671	38.54	9.671	4.100	263
Qt 18	18.428	15.718	38.64	9.715	4.050	307
Qt 12	18.401	15.682	38.47	9.647	4.043	283
Q 71	18.448	15.693	38.48	9.664	4.018	263
Q 15	18.650	15.751	38.76	9.756	4.044	190
SL 4	18.455	15.723	38.61	9.722	4.083	294
CAD 1	18.614	15.693	38.66	9.646	4.008	144
AX 5	18.660	15.745	39.14	9.743	4.200	175
X 1	18.530	15.707	39.00	9.683	4.206	221
SVZ	18.342	15.654	38.43	9.598	4.051	291
Brittany—Cornwall						
Pc 5630	18.355	15.611	38.27	9.512	3.961	230
Pc 5615	18.256	15.594	38.15	9.489	3.960	280
Pc 5618	18.218	15.603	38.15	9.511	3.984	319
Pc 561	18.221	15.600	38.11	9.505	3.963	313
Brt 17	18.225	15.619	38.28	9.543	4.042	330
Brt 5	18.212	15.621	38.30	9.548	4.058	344
Fla	18.301	15.591	38.13	9.478	3.925	244
WR	18.399	15.624	38.33	9.532	3.966	214
RC 4	19.786	15.690	38.30			
Vosges—Black Forest						
STA	18.093	15.565	38.03	9.451	3.994	363
STB	18.141	15.558	38.00	9.431	3.950	320
Hck	18.125	15.573	38.05	9.463	3.984	349
CDF	18.039	15.553	38.01	9.434	4.010	387
SAUL	18.155	15.612	38.21	9.537	4.051	374
FN 3	18.095	15.585	38.06	9.491	4.010	385
FN 12	18.200	15.615	38.25	9.538	4.041	346
FN 1	18.025	15.583	38.08	9.496	4.060	433
FN 13	18.140	15.601	38.12	9.517	4.015	372
Morocco						
MII C 42	18.075	15.601	38.15	9.525	4.068	418
M A 42	18.081	15.610	38.19	9.543	4.085	425
French Massif Central						
Lz 3	18.393	15.658	38.37	9.600	3.996	260
Lz 12	18.339	15.644	38.30	9.579	3.989	280
Lz 8	18.398	15.660	38.47	9.604	4.038	259
MLV	18.337	15.683	38.48	9.657	4.080	329
Marg	18.385	15.695	38.55	9.675	4.087	310
Peyr	18.245	15.638	38.37	9.578	4.077	341
1189	18.193	15.653	38.33	9.615	4.091	396
1356	18.212	15.640	38.30	9.592	4.065	371
1355	18.258	15.688	38.60	9.676	4.184	391
1354	18.246	15.654	38.36	9.610	4.072	359
Br 1	18.445	15.724	38.59	9.725	4.077	302
SIDO	18.383	15.684	38.45	9.653	4.043	298
Iberian Peninsula						
RUG	18.253	15.623	38.26	9.547	4.016	318
FIN	18.181	15.616	38.32	9.542	4.086	361
RUF	18.147	15.609	38.23	9.532	4.062	377
PIN	18.565	15.643	38.27	9.553	3.854	117
GC 37	18.230	15.623	38.23	9.552	4.019	334
M 181	18.265	15.619	38.14	9.538	3.957	304
M 317	17.600	15.483	37.45	9.345	3.991	620
M 379	18.174	15.587	38.13	9.485	3.996	331
M 66	18.293	15.630	38.30	9.556	4.013	297
M 254	18.254	15.641	38.24	9.583	4.014	338
M 247	18.268	15.630	38.22	9.559	3.994	315
Corsica	18.300	15.628	38.32	9.552	4.019	290
Appalachians						
Mil 1	18.370	15.640	38.28	9.567	3.962	254
Mil 2	18.372	15.633	38.28	9.553	3.959	244
1372	18.344	15.631	38.18	9.552	3.930	262
GA 4B	18.345	15.590	38.04	9.471	3.860	211
Ly 3	18.486	15.634	38.39	9.543	3.945	163
1285	18.377	15.601	38.11	9.486	3.873	201
1302	18.385	15.607	38.14	9.500	3.883	203
1141	18.481	15.602	38.12	9.481	3.820	127
Granulitic xenoliths (whole rocks)						
2520	18.530	15.645	38.73			
803	18.190	15.663	38.49			
509	18.737	15.687	38.75			
802	18.535	15.693	39.10			
199	18.637	15.676	38.58			
168	18.266	15.670	38.50			
BAL 1	18.697	15.727	39.35			

The single stage model, μ, K and model age are calculated from the modern lead values given by Stacey and Kramers[16] and initial lead values of Tatsumoto *et al.*[17]. The analytical technique for lead separation is described in ref. 58 and the leaching procedure in ref. 56. Analytical uncertainties are reproducible from one measurement or another and are in the 1‰ range.

Provenance of granite samples:

Pyrénées—Ml, Maladetta; Qt, Q, Querigut; SL, Saint Laurent; CAD, Cady; AX, X, Ax les Thermes; SVZ, Salvezines.

Brittany—Pc, Ploumanach; Brt, south Brittany; Fla, Flamanville.

Cornwall—Wr, saint Austell; RC, Tregonning.

Vosges, Black Forest—STA, STB, Remirmont; HCK, Honeck; CDF, Champ du Feu; SAUL, Saulieu; FN, Black Forest.

Morocco—M 11, MA, Zeida.

Massif Central—Lz, Mt Lozere; MLV, Millevache; Marg, Margeride; Pey, Peyrol; 1189, St Mathieu; 1356, Brâme; 1355, 1354, St Sylvestre; Br, Brousse; SIDO, Sidobre.

Iberian Peninsula—RUG, FIN, RUF, PIN, Galicia; GC, Guadarrama; M, Sierra Morena.

Corsica—Vizavonna.

Appalachians—Mil 1 and Mil 2, Milford (NH); 1372, York (SC); GA 4B, Sparta (GA); Ly 3, Lyman (ME); 1285, Lilesville (NC); 1302, Pageland (SC); 1141, Castalia (NC).

All granulitic xenoliths come from Bournac (Massif Central).

roughly correlate with high U/Pb ratios (Table 2). For most of the 'normal' samples the measured $^{238}U/^{204}Pb$ values are <1, and therefore, for a given region, the least radiogenic samples should be close to the initial lead isotope composition of the parent magmas.

Three samples exhibit anomalously high $^{207}Pb/^{204}Pb$ values when compared with their expected regional value. All of these samples, Qt-15 (Querigut, Pyrénées), AX5 (Ax les Thermes, Pyrénées) and 1355 (Saint Sylvestre, Western Massif Central), present strong evidence of deuteric effects. Although the cause of the observed shift is unknown, it is probably related to country-rock interaction or to inherited U-rich mineral leaching. When applied to these samples, the single stage model using the modern lead isotopic composition of Stacey and Kramers[16] gives an age in the range 360–260 Myr which is in reasonable agreement with their actual age. Therefore we used this model together with the primordial lead values of Tatsumoto *et al.*[17] to estimate the μ values of the reservoir for each sample.

The apparent μ values vary from 9.45 (Vosges) to 9.70 (Pyrénées) and straddle the μ value (9.52) calculated for the so-called 'normal growth curve', whose parameters are used here.

Apparent $K(^{232}Th/^{238}U)$ values range from 3.94 to 4.10 (discarding the samples with anomalous $^{206}Pb/^{204}Pb$ ratios) but show no obvious geographical pattern.

The systematic variation of lead isotope composition across the Hercynian orogeny is shown in Figs 2 and 3. A plot of the apparent μ values on a map (Fig. 3) with the Iberian Peninsula

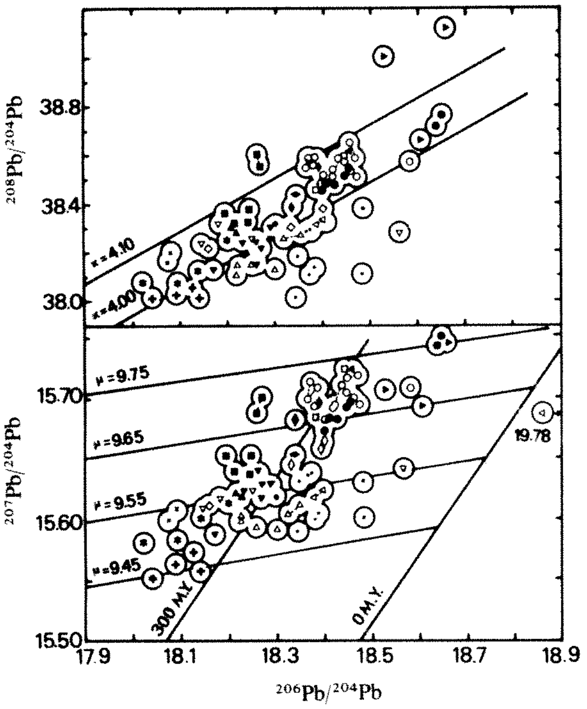


Fig. 1 Analytical data for Western Europe and eastern North America plutonic feldspars. Single stage evolution curves and isochrons are calculated from data in refs 16, 17. Different symbols indicate different geographical areas.

and Corsica in their predrift position[18,19], illustrates two fundamental points: a pattern of concentric zones centred on southern France with values decreasing outwards and apparently monotonous and continuous variations of the μ values. Palaeogeographical boundaries inferred for the Devonian[19] are also shown on Fig. 3 and their agreement with the zones is remarkable. Moreover, the position of the palaeogeographical domains did not change very significantly during most of the Palaeozoic.

The least radiogenic lead isotope compositions are located on the 'outer' (northern or western) edges of the orogeny (zone 3 of Fig. 3) where prominent Palaeozoic volcanism (spilites, dolerites and dominant keratophyre/rhyolites) is present[20,21]. In contrast, the magmatic and tectonometamorphic 'crest' running from Galicia through southern Brittany and Northern Massif Central to Central Europe[22] does not show in the lead isotope distribution. It has been proposed[5,23,24] that the prominent Hercynian slip fault running along this line could be the remnant of an intercontinental suture through which oceanic lithosphere (whose width has been estimated[22] to ~2,000 km) has been consumed during Lower and Middle Palaeozoic. Indeed, the only three known occurrences of Hercynian blue-amphibole have been reported along this line as well as rocks in the eclogite facies and true Hercynian peridotites. The present data do not support a discontinuity in the lead isotope composition across this boundary, though allowance must be made for the scale of our sampling and the effect of analytical uncertainties which may blur the observed variations. The very similar lead isotope compositions in granitic feldspars from Brittany and Galicia add another piece of evidence for the post-Hercynian opening of the Bay of Biscay[25].

We have also analysed a few samples from the eastern US. The K-feldspars from Hercynian rocks there have distinctive lead isotope compositions: $^{206}Pb/^{204}Pb$, 18.34–18.48; $^{207}Pb/^{204}Pb$, 15.59–15.64; $^{208}Pb/^{204}Pb$, 38.05–38.39 (Fig. 1). In the isotope diagram these compositions define a domain adjacent to, but significantly different from the domain for the European Hercynian granites. Both the $^{207}Pb/^{204}Pb$ ratios and the apparent μ values calculated for the source of granitic rocks decrease to the south-east when moving across the major geological boundaries, which suggests a pattern broadly similar to the one observed in Europe. Current strontium and oxygen isotope analyses of Hercynian granites from the eastern US[26–30] show that despite much lower initial $^{87}Sr/^{86}Sr$ ratios (0.703–0.706), these rocks have kept the $\delta^{18}O$ record of a previous alteration–sedimentation cycle. A crustal origin previously proposed for these rocks is confirmed by the continental signature of the lead isotopes. However, any attempt to rebuild a predrift model for the Hercynian orogeny, isotope data requires that the eastern US and Europe are distinctive segments of the Palaeozoic continental crust.

Genesis of Variscan crust according to Pb isotope data

On a $^{206}Pb/^{204}Pb$–$^{207}Pb/^{204}Pb$ diagram, the Variscan granitoids fall into the future (J) lead domain relative to single stage model calculated with the initial lead data of Tatsumoto *et al.*[17] and the U decay constants of Jaffey *et al.*[31]. Figure 4 shows the Hercynian granitoid domain (shifted 300 Myr forwards) and that for Mid-Ocean-Ridge basalts. It is impossible to make these two domains coincide. The model generating granitoids through partial melting of a subducting plate or through melting of mantle-type igneous rocks accumulated in an interarc basin, does not apply to Variscan granitoids from Europe and the eastern US. Even considering a period of the order of 10^9 yr, between the time of formation of the basalt series and that of their melting, it is impossible to generate the Hercynian granitoid domain, the apparent μ of the latter being much too high. However, the Hercynian granitoid domain (with isotopic compositions corrected for decay that would have occurred in source rocks during 300 Myr) compares well with high $^{207}Pb/^{204}Pb$ samples from the Caribbean Islands (ref. 32 and B. Dupre, unpublished results), while being markedly different to results from other areas, such as the Cascades[33], the Mariana Islands[34] or the Andes[35].

To explain the lead isotopic composition of Variscan granitoids, we may consider two extreme models on the basis of geological and geochemical observations.

(1) The crust–mantle mixing model: The granitoids are derived from a mixture of mantle and recycled preexisting continental crust[36,37].

The present lead isotopic data for Hercynian granitoids are compatible with this model. Values for lead of oceanic origin are variable[38,39] and the continental crust is heterogeneous. We may thus expect that a mixture of these two heterogeneous components could yield the observed pattern. However, as far as lead isotopes are concerned, the mantle can be of the oceanic ridge type but is unlike the source of ocean-island basalts or of continental flood-basalts. Note that this model is not based on a mixture of one crustal component and one mantle-derived component but on a mixture on an unspecified number of components of each type.

Table 2 Concentrations of Pb and U (ID) as determined by isotope dilution, U and Th by neutron activation

Sample	Pb (p.p.m.)	U (p.p.m.)	Th (p.p.m.)	$^{238}U/^{204}Pb$	$^{232}Th/^{238}U$
FIN	69.5	0.41 0.427 (ID)	1.52	0.5	3.70
PIN	33.8	5.57 5.52 (ID)	2.93	14	0.52
RUG	52.5	1.55 1.762 (ID)	6.11	2.6	3.94
CdF	24.6	0.20 0.196 (ID)	1.65	0.6	8.25
Qt 15	42.7	0.33	0.85	0.7	2.57
Qt 71	46.15	0.41	1.25	0.7	3.08
Ly 3	52.8	0.34	0.84	0.5	2.47
GA 4B	31.6	0.23	0.77	0.6	3.34
WR	60.24	0.23 0.203 (ID)	0.57	0.3	2.47
RC4	7.7	0.84 0.97 (ID)	0.137	11	0.16

J. L. Joron analyst, Laboratoire Pierre Süe.

(2) The crust-derived granitoid model: Here, granitoids are derived from a U- and Pb-zoned continental crust. Direct geochemical evidence[40], heat flow studies[41,42] as well as earlier isotopic studies[43–45] have suggested that the lower crust is depleted in U relative to the upper crust and should be characterized by significantly lower μ values. This model implies that granites with low μ values would derive from the lower crust and those with higher μ values from the intermediate crust. As a corollary to this model, ancient source rocks are indicated by the average slope of the ^{207}Pb/^{204}Pb, ^{206}Pb/^{204}Pb pattern (Fig. 3) (about 3,800 Myr for a secondary isochron). Even if such a figure does not accurately reflect the local primitive differentiation age of the crust—a point which could be really difficult to confirm—the initiation of these μ differences must be ancient (>3,000 Myr) to produce the observed ^{207}Pb/^{204}Pb variations.

Before discussing these models, we must point out that they are neither mutually exclusive nor at variance with more complex models which involve continuous or episodic mixing and reworking cycles[46,47]. The contamination discussed in model (1) can take place at different levels of the crust and thus the contaminating component may have highly variable μ values including very low ones. The problem is thus to determine when and where the crust–mantle mixing occurred. The following arguments seem to favour a crustal origin rather than a crust–mantle mixing in Hercynian time.

(1) The measured isotopic compositions for neodymium[37], strontium[48–52] and oxygen[53] in similar granitoids indicate that the granite source is the continental crust and not the mantle.

(2) We have measured the lead isotopic composition for seven granulite-facies xenoliths from Bournac, France, for which preliminary data on zircons suggest an age of ~300 Myr (Table 1). These specimens are a representative sample of the catazone for the French Massif Central[54,55]. No accurate correction for uranium decay during the past 300 Myr can be applied to the data due to the lack of U and Pb concentrations for these particular samples, so their field is shown in Fig. 4 without correction. Nevertheless, the assumption of μ values in a range reasonable for this rock type (0–5), can obviously make the domains of granulitic xenolith and Hercynian granitic rocks from the same area coincide, thus permitting a genetic link between them. Although this coincidence is only shown for the Massif Central we suggest that it can be extended to areas of distinct ^{207}Pb/^{204}Pb values.

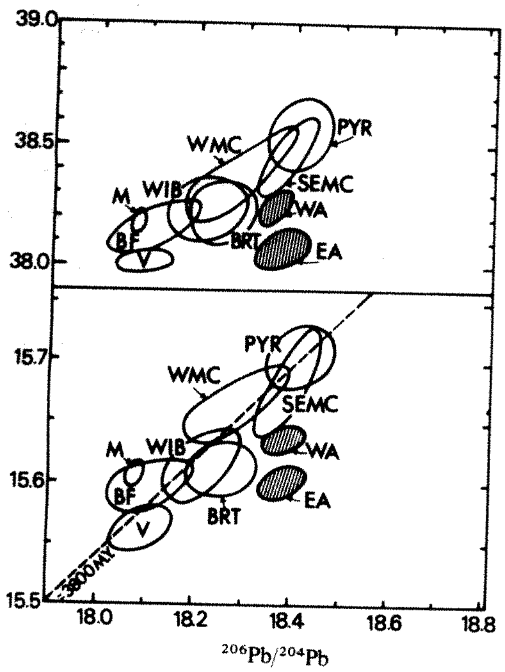


Fig. 2 Distribution of data according to their geographical origin A few anomalous radiogenic samples have been discarded. PYR, Pyrénées; WMC, western Massif Central; SEMC, southeastern Massif Central; WIB, west of the Iberian Peninsula; BRT, Brittany and Cornwall; M, Morocco; BF, Black Forest; V, Vosges. Hatched fields: WA, Western Appalachians; EA, Eastern Appalachians.

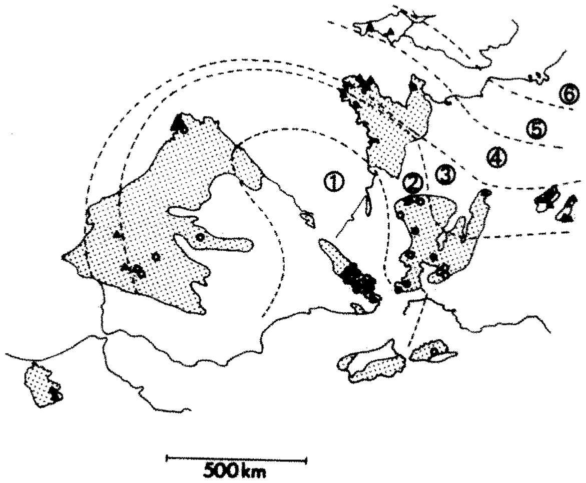


Fig. 3 Geographical plot of single stage μ values for plutonic feldspar in Western Europe on a predrift reconstruction proposed by Le Pichon[18]. Approximate palaeogeographical boundaries (dashed line) and palaeogeographic zones numbered 1 to 6 (6 is the stable Europe) are simplified from Matte[19]. Model μ values: ●, >9.65; ○, 9.55–9.65; ▲, 9.45–9.55; ◆, <9.45. □, Prehercynian outcrops.

(3) The very existence of isotopic provinces is better understood in such a model as these provinces are likely to reflect variations in orogenic segment characteristics (thickness, chemical and petrological compositions, age). In this way, the isotopic characteristics are expected to parallel the geological provinces defined from structural observations and metamorphic petrology[19].

(4) In the crust–mantle mixing model, we could anticipate that on a regional scale, Vosges or Cornwall granitoids should present a much more pronounced mantellic signature than those from the southern France. Petrographic and geochemical evidence for such a distinction is virtually absent.

Consequences for the orogenic processes

The present results suggest that Variscan granitoids were generated principally (if not totally) by remobilizing the previously differentiated continental crust. In western Europe and in the eastern US, at least, no new segment of continental crust has been formed during the Hercynian event. Such a result agrees with recent Sm–Nd and Rb–Sr data[37] according to which very little continental crust has been formed in recent time and the young crust results from a mere recycling of preexisting

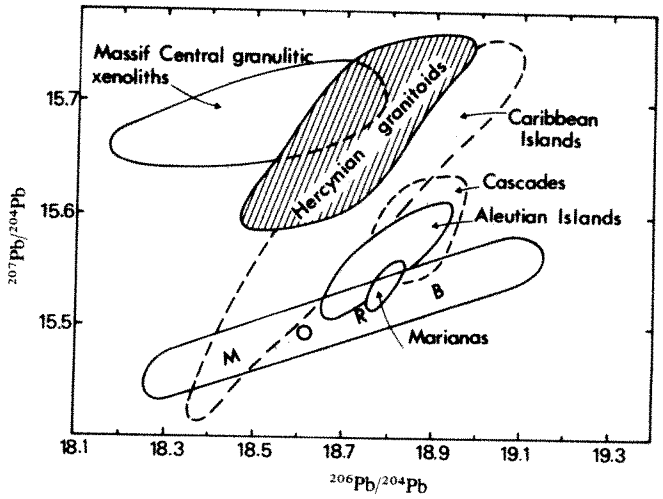


Fig. 4 ^{207}Pb/^{204}Pb versus ^{206}Pb/^{204}Pb fields of some present day volcanic suites : MORB (Mid-Ocean-Ridge Basalt)[59], Marianas[34], Cascades[33], Caribbean (ref 32 and B. Dupre, unpublished results), Aleutian[60], compared with the field of Hercynian plutonic feldspar data shifted 300 Myr forwards in time and to that of granulitic xenoliths from Massif Central.

continental material, whereas the interpretation of lead and strontium isotope data may differ for the Precambrian.

A second point concerns the behaviour of lead isotopes during the genesis of granitoid rocks. Although strontium and oxygen isotopic compositions may vary even within a single pluton[56], lead isotopes are remarkably uniform on a regional scale. As observed for the Maladeta intrusion[56], and at Ploumanac'h (Brittany), the lead isotopic ratios are virtually indistinguishable among different rock types from a single composite pluton (gabbro, granodiorite, granite). In our opinion, this requires their being levelled at the local average for the crustal lead by fairly efficient transport and/or reequilibration mechanisms. This observation implies that a large scale circulation of fluid does exist during such processes, remobilizing and homogeneizing lead isotopes, much more than strontium or oxygen.

Such an important circulation of fluids during melting of the lower crust may be responsible for the zoning of radioactive elements[57]. It may also induce remobilization and enrichment in incompatible elements such as Pb, Zn, Mo, Ta, Sn, U of the continental crust. The orogenic remobilization may be the origin of ore deposits associated with granitoids thereby yielding the so-called metallogenic provinces. When distinctive regional chemical characteristics exist, they are perpetuated through geological time through recycling and fluid remobilization processes. In the present view, lead isotopic provinces and metallogenic provinces will coincide as both result from a pervasive fluid circulation within the Earth's crust.

We thank the following colleagues for supplying samples: B. Azambre, M. Brunel, R. Capdevilla, B. Charoy, M. Dagallier, J. L. Duthou, S. Fourcade, F. D. Fullagar, H. E. Gaudette, J. Hamet, G. Ibarguchi, J. Lameyre, J. Leterrier, also L. T. Silver and B. Dupre for fruitful discussions, and A. Juery for assistance with some of the measurements. Part of this work was supported through CNRS ATP-39 82 and, in part, through a NATO fellowship while A.M.V. and F.A. were visiting at Caltech. This is IPGP contribution NS/500.

Received 27 November 1980; accepted 2 April 1981.

1. Zwart, H. J. *Geol. Mijnb.* **46,** 283–309 (1967).
2. Krebs, W. & Wachendorf, H. *Bull. geol. Soc. Am.* **84,** 2611–2630 (1973).
3. Nicolas, A. *Nature* **236,** 221–223 (1972).
4. Dewey, J. F. & Burke, K. C. A. *J. Geol.* **81,** 683–693 (1973).
5. Mattauer, M. 2*nd Reun. Ann. Sci. Terre Abstr.* **279,** (1974).
6. Hurley, P. M., Hugues, H., Faure, G., Fairbairn, H. W. & Pinson, W. A. *J. geophys. Res.* **67,** 5315–5334 (1962).
7. Patterson, C. & Tatsumoto, M. *Geochim. cosmochim. Acta* **28,** 1–22 (1964).
8. Doe, B. R. *J. Petrol.* **8,** 51–83 (1967).
9. Kistler, R. W. & Peterman, Z. E. *Bull. geol. Soc. Am.* **84,** 3489–3512 (1973).
10. Armstrong, R. L., Taubeneck, W. H. & Hales, P. O. *Bull. geol. Soc. Am.* **38,** 397–411 (1977).
11. Taylor, H. P. & Silver, L. T. *Int. Conf. Geochron. Cosmochron. Isotope Geol.* Aspen, 423–426 (1978).
12. Moorbath, S. & Welke, H. *Earth planet. Sci. Lett.* **5,** 217–230 (1968).
13. Zartman, R. E. *Econ. Geol.* **69,** 792–805 (1974).
14. Rogers, J. J. W., Adams, J. A. S. & Galting, B. *Am. J. Sci.* **263,** 817–822 (1965).
15. Rosholt, J. N., Zartmann, R. E. & Nkomo, J. T. *Bull. geol. Soc. Am.* **84,** 989–1002 (1973).
16. Stacey, J. S. & Kramers, J. D. *Earth planet. Sci. Lett.* **26,** 207–221 (1975).
17. Tatsumoto, M., Knight, T. J. & Allègre, C. J. *Science* **180,** 1279–1283 (1973).
18. Le Pichon, X. *Tectonophysics* **38,** 169–209 (1977).
19. Matte, P. *Nova Acta Leopoldina* **45,** 239–262 (1976).
20. Bebien, J., Rocci, G., Floyd, P. A., Juteau, Th. & Sagon, J. P. *Coll. Interface* (CNRS, Rennes, 1974).
21. Boyer, C. *Bull. Bur. Rech. geol. Miner., Orléans* Section No. 1, 3–24 (1976).
22. Bard, J. P., Burg, J. P., Matte, P. & Ribeiro, A. *Proc. 26th int. Congr. Geology*, Paris C**6,** 233–246 (1980).
23. Burg, J. P. & Matte, Ph. *Z. Geol. Ges.* **129,** 429–460 (1978).
24. Montigny, R. & Allègre, C. J. *C.r. hebd. Séanc. Acad. Sci., Paris* **279**D, 543–545 (1974).
25. Choukroune, P., Le Pichon, X., Seguret, M. & Sibuet, J. C. *Earth planet. Sci. Lett.* **18,** 109–118 (1973).
26. Fullagar, P. D. *Bull. geol. Soc. Am.* **82,** 2845–2863 (1971).
27. Fullagar, P. D. & Buttler, J. R. *Bull. geol. Soc. Am.* **8,** 409–416 (1976).
28. Fullagar, P. D. & Buttler, J. R. *Am. J. Sci.* **279,** 161–185 (1979).
29. Wenner, D. B., Whitney, J. A. & Stormer, J. C. Jr *Geol. Soc. Am. Progr. Abstr.* 1221 (1977).
30. Sinha, K. A. & Merz, B. *Geol. Soc. Am. Progr. Abstr.* 493 (1978).
31. Jaffey, A. H., Flynn, K. F., Glendenin, L. E., Bentley, W. C. & Essling, A. M. *Phys. Rev.* **C4,** 1889–1906 (1971).
32. Armstrong, R. L. & Cooper, J. A. *Bull. volcan.* **35,** 27–63 (1971).
33. Church, S. E. *Earth planet. Sci. Lett.* **29,** 175–188 (1976).
34. Meijer, A. *Bull. geol. Soc. Am.* **87,** 1358–1369 (1976).
35. Tilton, G. R. & Barreiro, B. A. *Science* **210,** 1245–1247 (1980).
36. Allegre, C. J. & Dars, R. *Geol. Rdsch.* **55,** 226–237 (1965).
37. Allegre, C. J. & Ben Othman, D. *Nature* **286,** 335–342 (1980).
38. Sun, S. S. & Hanson, G. N. *Geology* **3,** 297–302 (1975).
39. Dupre, B. & Allègre, C. J. *Nature* **286,** 17–22 (1980).
40. Heier, K. S. & Adams, J. A. S. *Geochim. cosmochim. Acta* **29,** 53–61 (1965).
41. Roy, R. F., Blackwell, D. D. & Birch, F. *Earth planet. Sci. Lett.* **5,** 1–12 (1968).
42. Lachenbruch, A. H. *J. geophys. Res.* **75,** 3291–3300 (1970).
43. Doe, B. R., Tilling, R. I., Hedge, C. E. & Klepper, M. R. *Econ. Geol.* **63,** 884–906 (1968).
44. Zartman, R. E. & Wasserburg, G. J. *Geochim. cosmochim. Acta* **33,** 901–942 (1969).
45. Gray, C. M. & Oversby, V. M. *Geochim. cosmochim. Acta* **36,** 939–952 (1972).
46. Armstrong, R. L. *Rev. Geophys. Space Phys.* **6,** 175–199 (1968).
47. Doe, B. R & Zartman, R. E. in *Geochemistry of Hydrothermal Ore Deposits* (ed. Barnes, H. L.) 22–70 (Wiley, New York, 1979).
48. Vitrac-Michard, A. & Allègre, C. J. *Contr. Miner. Petrol.* **50,** 257–285 (1975).
49. Hamet, J. & Allègre, C. J. *Bull. geol. Soc. Am.* **87,** 1429–1442 (1976).
50. Vidal, P. thesis, Univ. Rennes (1976).
51. Duthou, J. L. thesis, Univ. Clermont-Ferrand (1977).
52. Peucat, J. J., Charlot, R., Mifdal, A., Chantraine, J. & Autran, A. *Bull. Bur. Rech. geol. Miner., Orléans* 349–356 (1979).
53. Albarede, F. & Fourcade, S. 26*th Congr. int. Geology* Paris, Abstr. V.II 755 (1980).
54. Leyreloup, A. thesis, Univ. Nantes (1973).
55. Leyreloup, A., Dupuy, C. & Andriabololona, R. *Contr. Miner. Petrol.* **62,** 283–300 (1977).
56. Michard-Vitrac, A., Albarede, F., Dupuis, C. & Taylor, H. P. Jr *Contr. Miner. Petrol.* **72,** 57–72 (1980).
57. Albarede, F. *Earth planet. Sci. Lett.* **27,** 73–78 (1975).
58. Manhes, G., Minster, J. F. & Allègre, C. J. *Earth planet. Sci. Lett.* **39,** 14–24 (1978).
59. Tatsumoto, M. *Earth planet. Sci. Lett.* **38,** 63–87 (1978).
60. Kay, R. W., Sun, S. S. & Lee-Hu, C. N. *Geochim. cosmochim. Acta* **42,** 263–273 (1978).

Only one of two closely related yeast suppressor tRNA genes contains an intervening sequence

M. V. Olson*, G. S. Page†, A. Sentenac‡, P. W. Piper§, M. Worthington, R. B. Weiss & B. D. Hall

Department of Genetics, University of Washington, Seattle, Washington 98195, USA
* Department of Genetics, Box 8031, Washington University School of Medicine, St Louis, Missouri 63110, USA
† Departments of Biochemistry and Biophysics, University of California, San Francisco, California 94113, USA
‡ Service Biochimie, Départment de Biologie, Centre d'Etudes Nucléaires de Saclay, BP No. 2, 91190, Gif-sur-Yvette, France
§ Department of Biochemistry, University College London, Gower Street, London WC1E 6BT, UK

The yeast genes that code for the serine-inserting SUP-RL1 amber *and* SUQ5 ochre *suppressors have been cloned and sequenced. These two unlinked genes differ by only three base pairs in their coding regions yet they encode tRNAs of different translational specificities, and while the* SUP-RL1 *gene has a 19-base pair intervening sequence, the* SUQ5 *gene has none.*

SINCE the selection of yeast nonsense suppressors was first described by Hawthorne and Mortimer[1], the study of these genes has contributed to several significant advances in eukaryotic molecular genetics. Yeast suppressor strains, for example, were used to prove that eukaryotes and prokaryotes use the same nonsense codons[2,3]; they were the source of transfer RNA in the first demonstrations of *in vitro* suppression in eukaryotic translation systems[4,5]; they were the first eukaryotic mutants to be sequenced both at the levels of RNA[6] and DNA[7], and they played a part in the discovery of intervening sequences and RNA splicing[7–10].

Most work on yeast nonsense suppressors has involved the tyrosine and serine-inserting suppressor genes. The tyrosine-inserting suppressors derive from a set of eight unlinked tyrosine

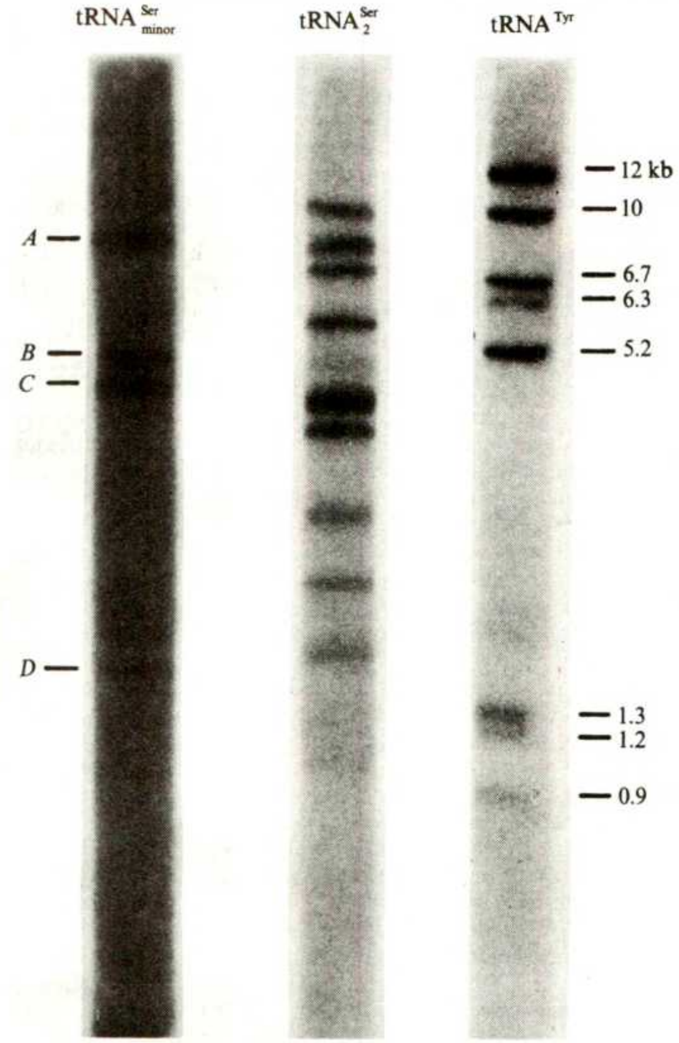


Fig. 1 Hybridization of three different ^{125}I-labelled yeast tRNA probes to *Eco*RI-cleaved, gel-fractioned yeast DNA. Yeast DNA from the strain Y4A was cleaved with *Eco*RI and fractionated on an agarose gel before being transferred to nitrocellulose[49]. Preparation of yeast DNA, electrophoresis, labelling of the tRNAs and hybridization were all as previously described[11,13]. The tRNA$^{Ser}_{minor}$ mixture was fractionated on and eluted from a two-dimensional acrylamide gel[15,23]. The purification of tRNA^{Tyr} and tRNA$^{Ser}_{2}$ has also been described elsewhere[11,34]. The sizes of the four tRNA$^{Ser}_{minor}$-hybridizing fragments are ~8.3, 4.7, 4.3 and 1.5 kilobases (kb).

tRNA genes, all of which code for the same mature tRNA. Reports are available on the cloning and sequencing of both wild-type and mutant tRNA^{Tyr} genes[7,11,12] as well as on the cross-correlation of the physical and genetic descriptions of the entire eight gene family[13].

The yeast serine-inserting suppressors derive from a family of tRNA genes that is both less studied and more complex than the tyrosine case. There are six serine codons (UCC, UCU, UCA, UCG, AGU and AGC), which are translated by at least four different classes of serine tRNAs[14–17]. Genetic analysis has revealed three different loci that can mutate to serine-inserting nonsense suppressors: the recessive lethal *amber* suppressor *SUP-RL1* (chromosome III)[8,18] and the *ochre* suppressors *SUQ5* (chromosome XVI)[19–22] and *SUP17* (chromosome IX)[19].

SUP-RL1 was shown by Piper[15] to cause the alteration CGA→CUA in the anticodon of a unique component of a heterogeneous set of minor serine tRNAs, which we designate tRNA$^{Ser}_{minor}$. It was recently demonstrated that the *sup-RL1*$^{+}$ gene product is the only component of tRNA$^{Ser}_{minor}$ whose anticodon is complementary to UCG, and a precursor to this tRNA containing a 19-base pair intervening sequence was also characterized[16].

We have performed a molecular analysis of the serine-inserting suppressor genes starting with a hybridization probe made by labelling the tRNA$^{Ser}_{minor}$ mixture. Here we demonstrate that the tRNA$^{Ser}_{minor}$ species are transcribed from a family of four genes, all of which have been cloned from a wild-type yeast strain, and two of which have been identified with the genetic loci *SUP-RL1* and *SUQ5*. DNA sequences have been determined for wild-type and suppressor alleles of both genes. One striking feature of the results is that although the *SUP-RL1* and *SUQ5* genes differ by only three base pairs within their coding regions, the *SUP-RL1* gene has an intervening sequence whereas the *SUQ5* gene does not.

The tRNA$^{ser}_{minor}$s are encoded by a four-gene family

To characterize and clone the genes for the various tRNA$^{Ser}_{minor}$ species, we used ^{125}I-labelled tRNA$^{Ser}_{minor}$ as an RNA–DNA hybridization probe. The tRNA$^{Ser}_{minor}$ mixture was purified by two-dimensional gel electrophoresis[23] and then iodinated by a procedure that specifically labels the modified nucleoside isopentenyl adenosine[11]. The presence of this nucleoside in all the components of tRNA$^{Ser}_{minor}$ offers a convenient means of labelling these RNAs efficiently and relatively specifically.

The number of genes in the tRNA$^{Ser}_{minor}$ family was estimated by examining the hybridization of ^{125}I-labelled tRNA$^{Ser}_{minor}$ to *Eco*RI-cleaved, size-fractionated yeast DNA. As shown in Fig. 1, there are four different yeast *Eco*RI fragments that hybridize to tRNA$^{Ser}_{minor}$. Comparison of the tRNA$^{Ser}_{minor}$ hybridization pattern with that produced by ^{125}I-labelled tRNA$^{Ser}_{2}$ indicates that at least three and probably all four of the tRNA$^{Ser}_{minor}$-hybridizing *Eco*RI fragments are distinct from those that hybridize to tRNA$^{Ser}_{2}$.

All four of the tRNA$^{Ser}_{minor}$-hybridizing *Eco*RI fragments were cloned in the λ insertion vector 607 (ref. 24). Plaques from a pool of λ–yeast recombinants, prepared by ligating together a mixture of *Eco*RI-cleaved λ and total yeast DNA, were screened using ^{125}I-labelled tRNA$^{Ser}_{minor}$ as the hybridization probe. Recombinant λ were isolated containing yeast *Eco*RI fragments that were identical in size to each of the four fragments *A–D*. Both the recombinant λ plaques and the cloned *Eco*RI fragments themselves showed strong hybridization to ^{125}I-labelled tRNA$^{Ser}_{minor}$.

Chromosomal location of two of the genes

Various RNA–DNA and DNA–DNA hybridization experiments were carried out to identify one or more of the tRNA$^{Ser}_{minor}$-hybridizing *Eco*RI fragments with known suppressor loci. We examined the hybridization of ^{125}I-labelled tRNA$^{Ser}_{minor}$ and some of the ^{32}P-labelled cloned *Eco*RI fragments to DNA from yeast strains that were aneuploid for chromosomes to which serine-inserting suppressor loci have been mapped. Strains that are disomic for a single chromosome ($n+1$) are particularly useful in this type of analysis. In the simplest case, one member of a set of bands produced by a particular probe may have enhanced intensity when DNA from a specific disomic strain is used. This behaviour implies that this band arises from hybridization to a restriction fragment that is derived from the disomic chromosome. Alternatively, a particular sequence that hybridizes to the probe may exist on restriction fragments of different sizes on the two homologous chromosomes present in the disome; in this case an extra band appears in hybridizations to disomic DNA. Both of these situations have been observed before[13,25] and were encountered again here.

The first disomic strain that was examined contained an extra copy of chromosome III, the chromosome to which the *SUP-RL1* locus maps. Figure 2*a* shows that, compared with a haploid control, an enhanced *C* band is observed when ^{125}I-labelled tRNA$^{Ser}_{minor}$ hybridizes with *Eco*RI-cleaved DNA from the chromosome III disome. This result establishes the tRNA$^{Ser}_{minor}$ gene on the *C* fragment as a candidate for the *SUP-RL1* gene. The same DNA filter that was used for the ^{125}I-labelled tRNA$^{Ser}_{minor}$ hybridization was subsequently used in hybridization with ^{32}P-labelled DNA from the cloned *C* fragment. Once

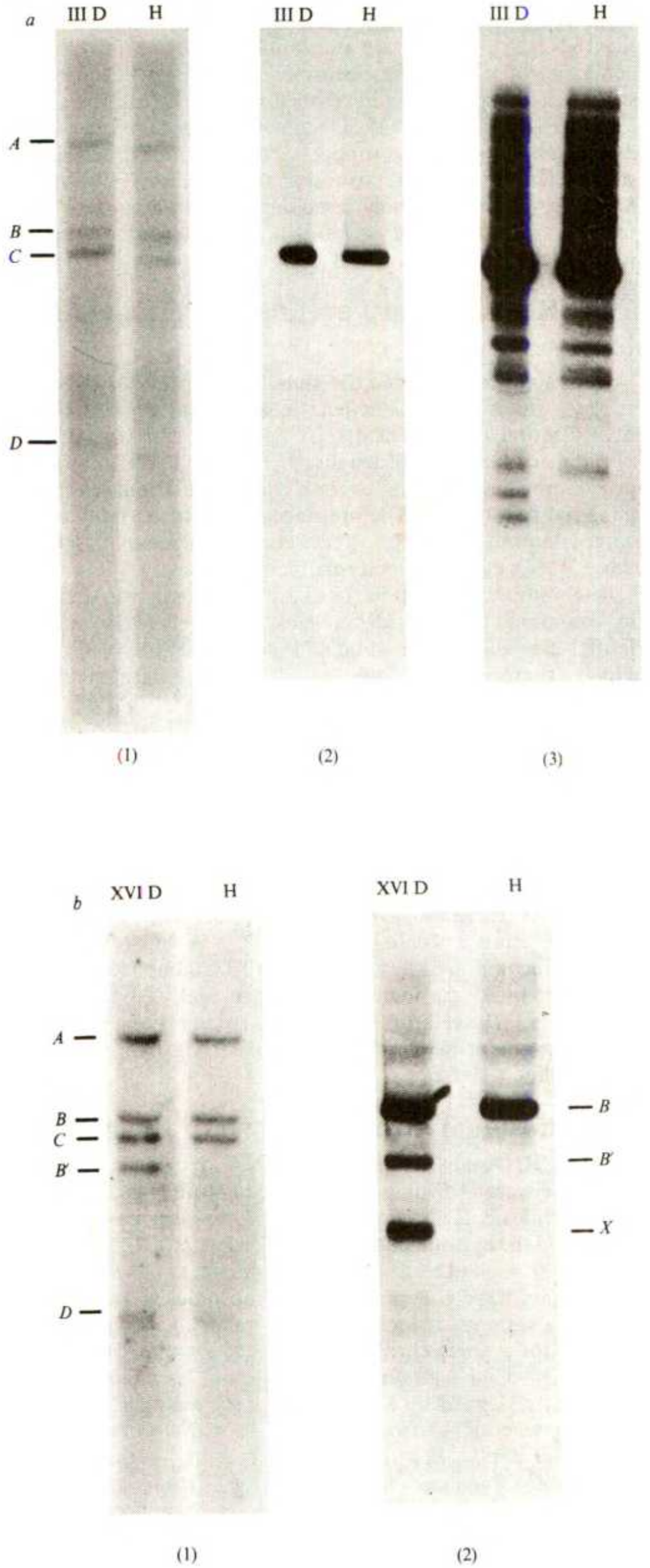


Fig. 2 Hybridization of ^{125}I-labelled tRNA$^{\text{Ser}}_{\text{minor}}$ and ^{32}P-labelled fragment-specific probes to DNA from aneuploid and haploid yeast strains. *a*, Comparison of the hybridization of ^{125}I-labelled tRNA$^{\text{Ser}}_{\text{minor}}$ (1) and ^{32}P-labelled cloned *C* fragment (2 and 3) with DNA from a chromosome 3 disome and a haploid yeast strain. The photograph in (3) is a longer exposure of the same autoradiogram as in (2). The haploid strain (96C) was derived from the disome (96CH) by mitotic chromosome loss. *b*, Comparison of the hybridization of ^{125}I-labelled tRNA$^{\text{Ser}}_{\text{minor}}$ (1) and ^{32}P-labelled cloned *B* fragment (2) with DNA from a chromosome XVI disome and a haploid yeast strain. The haploid strain (DH4093-5DH) was derived from the disome (DH4093-5D) by mitotic chromosome loss. (Both strains were provided by D. C. Hawthorne.) Experimental details were as described in Fig. 1 legend except that the hybridization conditions for the nick-translated, agarose gel-purified restriction fragments were as defined elsewhere[13].

again, an enhanced *C* band was observed with the chromosome III disome DNA. On longer exposure of the autoradiogram, it was also observed that the ^{32}P-labelled *C* fragment hybridizes quite strongly with a large number of other *Eco*RI fragments, presumably because of the presence on this fragment of one or more dispersed repetitive sequences.

Similar experiments using DNA from a chromosome XVI disome allowed a tentative identification of the tRNA$^{\text{Ser}}_{\text{minor}}$ gene on the *B* fragment with the serine-inserting suppressor locus *SUQ5*. In this case, however, instead of enhancement of the intensity of a pre-existing band, a new band appeared in the hybridization pattern of the disomic strain (Fig. 2*b*). This behaviour would be consistent with the possibility that one of the three unmapped *Eco*RI fragments (*A*, *B* or *D*) derives from chromosome XVI and that it exists in two variant forms on the disome's two copies of this chromosome. Because there was no *a priori* basis on which to associate the new band with a particular one of these three fragments, we performed hybridization experiments with the ^{32}P-labelled cloned *A*, *B* and *C* *Eco*RI fragments to identify the new band.

The first case tested, the *B* fragment, was found to have extensive homology with the disome's new tRNA$^{\text{Ser}}_{\text{minor}}$-hybridizing *Eco*RI fragment. When DNA from the disome was used, the ^{32}P-labelled *B* fragment hybridized strongly to three *Eco*RI fragments, labelled *B*, *B'* and *X* in Fig. 2*b*. *B'* is identical in size to the 'extra' ^{125}I-labelled tRNA$^{\text{Ser}}_{\text{minor}}$ fragment observed with disome DNA, while *X* is a fragment that does not hybridize to the tRNA. In the haploid, ^{32}P-labelled *B* hybridizes strongly only to the normal *B* fragment. These results indicate that in the disome one of the chromosome XVI homologues contains *B* fragment sequences on the two variant fragments *B'* and *X*, with the tRNA$^{\text{Ser}}_{\text{minor}}$ coding region lying on *B'*. The results in Fig. 2b also show that the cloned *B* fragment, like *C*, contains one or more dispersed repetitive elements.

Although the experiments on the chromosomal origins of *B* and *C* suggested the tentative assignments *B* = *SUQ5* and *C* = *SUP-RL1*, more direct evidence was required before unequivocal assignments could be made. Our next goal, therefore, was to determine the DNA sequences of the tRNA$^{\text{Ser}}_{\text{minor}}$ genes on *B* and *C*, as cloned from both wild-type and suppressor yeast strains. The details of the re-cloning of the *B* and *C* fragments from suppressor strains are described in Fig. 3 legend. Although the re-cloning experiments were basically straightforward, two minor points required special attention: first, the presence of dispersed repetitive DNA on both the *B* and *C* fragments complicated the choice of *B*- and *C*-specific DNA–DNA hybridization probes, and second, the fact that *SUP-RL1* can only be maintained in the heterozygous state required that potential *SUP-RL1*$_{\text{UAG}}$ clones be isolated from a λ-yeast recombinant phage pool that contained equal numbers of clones of the *SUP-RL1*$_{\text{UAG}}$ and *sup-RL1*$^{+}$ alleles. The first problem was solved by identifying subfragments of *B* and *C* that contain little or no dispersed repetitive DNA, and then using plasmids containing these predominantly single-copy subfragments as specific hybridization probes. Our approach to the second problem was to screen λ clones of the *C* fragment from a *SUP-RL1*$_{\text{UAG}}$/*sup-RL1*$^{+}$ heterozygote for a clone that was lacking the *Taq*I site that is present in the tRNA coding region of the wild-type gene. On the basis of the tRNA sequence, this site is predicted to be eliminated by the mutation of *sup-RL1*$^{+}$ to *SUP-RL1*$_{\text{UAG}}$.

One unexpected finding during the re-cloning experiments was that the *B* and *C* fragments were both a few hundred base pairs smaller in the suppressor strains than in our original wild-type strain. As shown in Fig. 3, the missing DNA in the suppressor strains in both cases maps to a region near the 5' end of the tRNA coding region. It should be emphasized that the suppressor strains were laboratory stocks that were not closely related to the original 'wild-type' strain. As a high degree of strain-to-strain variation in restriction fragment sizes has been previously documented for tRNA$^{\text{Tyr}}$-hybridizing restriction fragments[13], the mere presence of this variation among

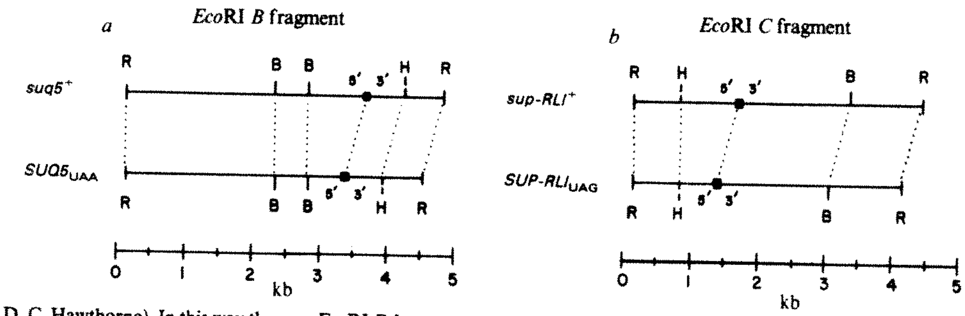


Fig. 3 Restriction maps of the tRNA$^{\mathrm{Ser}}_{\mathrm{minor}}$-hybridizing *Eco*RI fragments *B* and *C*. B, *Bam*HI; H, *Hin*dIII; R, *Eco*RI; ■ tRNA coding region. *a*, The map labelled *suq5*$^+$ refers to the original clone (λPM15) of the *Eco*RI *B* fragment from the wild-type yeast strain Y4A. The 1.4-kilobase (kb) *Bam*HI/*Hin*dIII subfragment was transferred to the plasmid vector pBR322 for fine structure restriction mapping and DNA sequencing. This plasmid (pPM15) was also nick-translated and used to screen a new λ607 pool of yeast *Eco*RI fragments, derived from the *SUQ5-ochre* strain 4093-4A (obtained from D. C. Hawthorne). In this way the new *Eco*RI *B* fragment, labelled *SUQ5*$_{\mathrm{UAA}}$ was obtained in a λ clone (λPM16). The 1.1-kilobase *Bam*HI/*Hin*dIII subfragment of the *SUQ5*$_{\mathrm{UAA}}$ *B* fragment was transferred to pBR322 to produce the plasmid pPM16, which was used in the DNA sequencing. *b*, The map labelled *sup-RL1*$^+$ refers to the original clone (λPM15) of the *Eco*RI *C* fragment from the wild-type strain Y4A. A strategy almost identical to that described for *a* was used to re-clone the *C* fragment from the *SUP-RL1 amber* strain DBD339 (obtained from D. Botstein). The tRNA$^{\mathrm{Ser}}_{\mathrm{minor}}$ gene on the *Eco*RI *C* fragment from Y4A was on a 2.5-kilobase *Bam*HI/*Hin*dIII subfragment, which was transferred to pBR322 to produce the plasmid pPM5—this was used in the DNA sequencing. A plasmid (pPM6) containing the 1.1-kilobase *Bam*Hi/*Eco*RI subfragment of the *Eco*RI *C* fragment was used as a hybridization probe to isolate several independent λ clones that contained DBD339 *Eco*RI *C* fragments. From these recombinants, we chose a clone (λPM10) that was missing the *Taq*I site found in the coding region of wild-type but not suppressor alleles of *SUP-RL1*. This clone, like all the DBD339 *C* fragment clones, contained a *C* variant that was 0.3 kilobases shorter than the Y4A *C* fragment. As shown in the map labelled *SUP-RL1*$_{\mathrm{UAG}}$, the 0.3-kilobase deficiency maps between the 5′ end of the tRNA$^{\mathrm{Ser}}_{\mathrm{minor}}$ coding region and the nearby *Hin*dIII site. The 2.2-kilobase *Bam*HI/*Hin*dIII subfragment that contains the tRNA$^{\mathrm{Ser}}_{\mathrm{minor}}$ coding region was transferred to pBR322 to produce the plasmid pPM11, which was used in the DNA sequencing.

tRNA$^{\mathrm{Ser}}$-hybridizing fragments is unsurprising. However, the similarity in the amount of the missing DNA and the position of the deficiency relative to the tRNA coding regions in the two tRNA$^{\mathrm{Ser}}_{\mathrm{minor}}$ genes is quite striking. This impression is reinforced in the case of the two *SUQ5* fragments by DNA sequencing data that locate one of the homology/non-homology transitions at position −20 relative to the start of the region that codes for mature tRNA (data not shown). The beginning of the tRNA gene and the site of the deficiency, therefore, map within the same 0.4% of the *Eco*RI *B* fragment. The size variation is not a cloning artefact because it is readily observable by DNA–DNA hybridization to size-fractionated restriction digests of the pertinent yeast DNAs using the cloned *B* and *C* fragments as probes (data not shown).

Identity of the *SUQ5* and *SUP-RL1* clones

The tRNA coding regions were precisely located by searching for the cluster of restriction sites that are predicted from the tRNA sequence to occur within tRNA$^{\mathrm{Ser}}_{\mathrm{minor}}$ coding regions. An end-label, partial digestion technique was used to visualize directly the clusters of sites[26,27]. DNA sequencing of all the wild-type and mutant clones was carried out by the method of Maxam and Gilbert[28], starting at restriction sites in or near the tRNA$^{\mathrm{Ser}}_{\mathrm{minor}}$ coding regions. The results are summarized in Fig. 4 and confirm the assignments *B* = *SUQ5* and *C* = *SUP-RL1*: the tRNA$^{\mathrm{Ser}}_{\mathrm{minor}}$ gene on *B*, when cloned from a *SUQ5*$_{\mathrm{UAA}}$ strain, codes for an *ochre* suppressor, and that on *C*, when cloned from a *SUP-RL1*$_{\mathrm{UAG}}$ strain, codes for an *amber* suppressor. In both cases the only changes observed in the coding regions of the two mutant genes are the expected alterations at the sites coding for the second nucleotide in the tRNA's anticodon. The results are also consistent with the conclusion from RNA sequencing that the tRNA$^{\mathrm{Ser}}_{\mathrm{minor}}$ mixture contains at least two functional classes of tRNA$^{\mathrm{Ser}}$ (ref. 16). The *sup-RL1*$^+$ gene codes for a tRNA whose anticodon is complementary to the serine codon UCG, while the *suq5*$^+$ gene codes for a tRNA whose anticodon is complementary to UCA.

Comparisons of the *sup-RL1*$^+$ and *suq5*$^+$ sequences reflects a close but unusual relationship between these two genes. The regions that code for mature tRNA differ at only three sites, including the functionally critical one that codes for the tRNA's wobble position base. The *sup-RL1*$^+$ gene, however, contains a

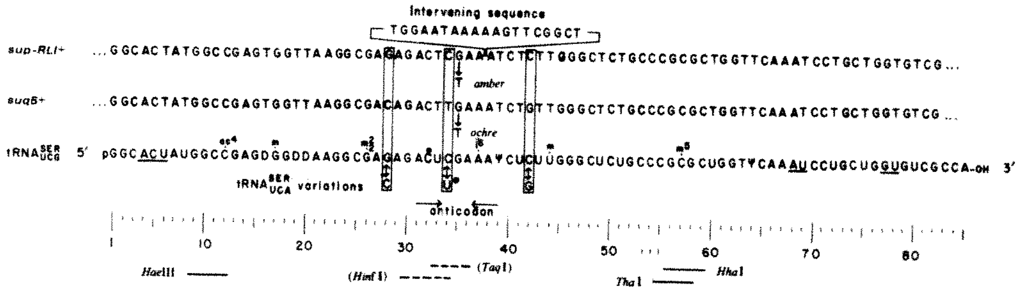

Intervening sequence
TGGAATAAAAAGTTCGGCT
sup-RL1+ ...GGCACTATGGCCGAGTGGTTAAGGCGAGAGACTCGAAATCTCTTGGGCTCTGCCCGCGCTGGTTCAAATCCTGCTGGTGTCG...
T *amber*
suq5+ ...GGCACTATGGCCGAGTGGTTAAGGCGACAGACTTGAAATCTGTTGGGCTCTGCCCGCGCTGGTTCAAATCCTGCTGGTGTCG...
T *ochre*
tRNA$^{\mathrm{SER}}_{\mathrm{UCG}}$ 5′ pGGCACUAUGGCCGAGDGGDDAAGGCGAGAGACUCGAAAΨCUCUUGGGCUCUGCCCGCGCUGGTΨCAAAUCCUGCUGGUGUCGCCA-OH 3′
tRNA$^{\mathrm{SER}}_{\mathrm{UCA}}$ variations
C
U*
G
anticodon
1
10
20
30
40
50
60
70
80
*Hae*III
(*Hinf*I)
(*Taq*I)
*Tha*I
*Hha*I

Fig. 4 The DNA sequences of the regions of the *SUP-RL1* and *SUQ5* genes that code for mature tRNA and the *SUP-RL1* intervening sequence. Only the (+)-strand is shown (that is, the strand whose sequence corresponds to that of the tRNA). The three sites of variation between the two wild-type genes are boxed. The two dotted restriction sites (*Hinf*I and *Taq*I) occur in the *sup-RL1*$^+$ but not the *suq5*$^+$ gene. The *Taq*I site is lost in the *SUP-RL1*$_{\mathrm{UAG}}$ allele. The tRNA sequence has been corrected so that it is consistent with the DNA sequence. Positions in the corrected tRNA sequence which differ from the published sequences[15,16] are underlined: (1) nucleotides 4–6 were misordered in the amino acid acceptor stem; (2) a complementary misordering occurred for nucleotides 77 and 78, which pair with nucleotides 4 and 5 in the mature tRNA structure; (3) the *SUP-RL1* gene is identical to that for *SUQ5* in the region that codes for positions 68 and 69, and the predicted RNA sequence is A-U. In previous work[15,16], a minor T1 oligonucleotide was observed that appeared to contain the sequence U-C in these positions. This 'U-C' oligonucleotide was present in molar amounts in the precursor to the *sup-RL1*$^-$ gene product, and the 'A-U' variant was absent from this precursor[16]. Consequently, the U-C sequence was assigned to tRNA$^{\mathrm{Ser}}_{\mathrm{UCG}}$ and the U-A sequence to the more abundant tRNA$^{\mathrm{Ser}}_{\mathrm{minor}}$ component, which we have now shown is partly a product of the *SUQ5* locus. The DNA sequence is not consistent with this interpretation and requires that both the *sup-RL1*$^+$ and *suq5*$^+$ gene products have the sequence AU. The C* at position 32 in the tRNA sequence is probably 3-methylcytosine (P.W.P., unpublished results), while the U* at position 34 is probably a derivative of 5-carboxymethyluridine (see text). The entire coding region sequence was determined for the *sup-RL1*$^+$ gene (plasmid pPM5) and the *SUP-RL1*$_{\mathrm{UAG}}$ allele (pPM11), and also for the *suq5*$^+$ gene (pPM15). In the case of the *SUQ5*$_{\mathrm{UAA}}$ allele, the sequence was determined only for positions +26 to +68. Considerable flanking sequence was also determined for three of the clones: *sup-RL1*$^+$ (pPM5), from −67 to +119; *SUP-RL1*$_{\mathrm{UAG}}$ (pPM11), from −73 to −10 and −1 to +97; *suq5*$^+$ (pPM15), from −100 to +119. Within the portions of the flanking regions where both pPM5 and pPM11 were sequenced, the sequences of the two clones were identical. Very little flanking sequence was determined for *SUQ5*$_{\mathrm{UAA}}$ (pPM16), but as discussed in the text, it was established that pPM15 and pPM16 are homologous from −7 to −20 and then diverge in sequence. The (+)-strand flanking sequences for *SUP-RL1* from −73 to +119, with asterisks at the site of the coding region, are as follows (slashes precede nucleotides −70, −60 . . . +90, etc.):

C(−73)TT/GGATCTGTGA/AAAAACGCCT/AGAAACATGT/ATAAAACCTA/CCGCTTTAAA/CCATAATTTC/AAATCGAAAT(−1)/
***T(+83)TTAATT/TTTTTTAAAT/AACATCGTTG/ATTAAAACAA(+119).

In the same format, the flanking sequences for *SUQ5* from −100 to +119 are as follows:

G(−100)TTATAAGGT/TGTTTCATAT/GTGTTTTATG/AACGTTCAGG/ATGACGTATT/GTCATACTGA/CATATCTCAT/TTTGAGATAC/
AACACTACCA/GATTATTTTG(−1)***T(+83)ATTATT/TTTTGAAATT/ATTTTTCAAT/AACTACCATG(+119).

Table 1 Relationships between the yeast $\mathrm{tRNA}^{\mathrm{Ser}}$ iso-acceptors that can translate the codons UCC, UCU, UCA and UCG

Iso-acceptor	Anticodon	Probable codon preferences	Apparent gene no.	Intervening sequence?	Corresponding suppressors
†$\mathrm{tRNA}^{\mathrm{Ser}}_{\mathrm{UCG}}$	CGA	UCG	1	Yes	*SUP-RL1*$_{\mathrm{UAG}}$
†$\mathrm{tRNA}^{\mathrm{Ser}}_{\mathrm{UCA}}$	U*GA	UCA	3	No	*SUQ5*$_{\mathrm{UAA}}$
†$\mathrm{tRNA}^{\mathrm{Ser}}_{2}$	IGA	UCA(?), UCC, UCU	~11	No	None known

In addition to our results, we used other data on $\mathrm{tRNA}^{\mathrm{Ser}}_{\mathrm{UCG}}$ (refs 15, 16, 18 and 50), $\mathrm{tRNA}^{\mathrm{Ser}}_{\mathrm{UCA}}$ (refs 15, 16) and on $\mathrm{tRNA}^{\mathrm{Ser}}_{2}$ (refs 14, 17 and 34). On the basis of sequencing studies of $\mathrm{tRNA}^{\mathrm{Ser}}$ species purified from brewer's yeast, a fourth iso-acceptor, $\mathrm{tRNA}^{\mathrm{Ser}}_{1}$, has been reported; it differs in sequence from $\mathrm{tRNA}^{\mathrm{Ser}}_{2}$ at three sites but has the same IGA anticodon[14]. This iso-acceptor is apparently absent from laboratory strains of *Saccharomyces cerevisiae*[15,17,34].

19-base pair intervening sequence but the *suq5*$^+$ gene has none. The *sup-RL1*$^+$ gene's intervening sequence, which was originally discovered by the RNA sequencing of a precursor to the mature tRNA[16], occurs at a uniquely defined position on the gene's transcript, one nucleotide away from the 3′ end of the anticodon. This positioning of the intervening sequence is highly, but perhaps not universally, conserved among tRNA genes[7,29–33].

The DNA sequences require minor revision of both published $\mathrm{tRNA}^{\mathrm{Ser}}_{\mathrm{minor}}$ sequences, as discussed in Fig. 4 legend. They also establish that the unidentified base present at position 32 in all the $\mathrm{tRNA}^{\mathrm{Ser}}_{\mathrm{minor}}$ species is a modified cytosine.

Outside their coding regions, the two genes display structural features that are common to many eukaryotic tRNA genes (see Fig. 4 legend). The CCA is not encoded, there is a presumed terminator tract of about 20 almost uninterrupted A·T pairs immediately after the coding region with a strong bias towards Ts in the (+)-strand and there is no obvious homology within the 5′ noncoding regions despite the near identity of the coding sequences.

Discussion

Our observations of the DNA sequences that code for components of the yeast $\mathrm{tRNA}^{\mathrm{Ser}}_{\mathrm{minor}}$ mixture, combined with the results of previous studies, lead to a schematic overview (Table 1) of the 'UCN' family of genes that code for serine tRNAs recognizing the UCC, UCU, UCA and UCG codons. Most of these genes code for $\mathrm{tRNA}^{\mathrm{Ser}}_{2}$ (ref. 34). This species also accounts for by far the major portion of the serine acceptor capacity of yeast tRNA[17]. The two genes whose sequences we report (*SUP-RL1* coding for $\mathrm{tRNA}^{\mathrm{Ser}}_{\mathrm{UCG}}$ and *SUQ5* coding for $\mathrm{tRNA}^{\mathrm{Ser}}_{\mathrm{UCA}}$) belong to a minor branch of the UCN family, which appears, on the basis of our RNA–DNA hybridization data, to contain four genes. The two sequenced cases differ from each other by only three nucleotides (4%) in the region coding for mature tRNA, while both differ by 11 nucleotides (13%) from the $\mathrm{tRNA}^{\mathrm{Ser}}_{2}$ genes. The 87% homology between $\mathrm{tRNA}^{\mathrm{Ser}}_{2}$ and the $\mathrm{tRNA}^{\mathrm{Ser}}_{\mathrm{minor}}$ species should not be taken as evidence of a particularly recent divergence between these two branches of the UCN family because yeast $\mathrm{tRNA}^{\mathrm{Ser}}_{2}$ even shares 78% homology with the rat liver $\mathrm{tRNA}^{\mathrm{Ser}}_{2}$ (ref. 35).

The special function of tRNAs with the U*GA anticodon (that is, the *suq5*$^+$ product $\mathrm{tRNA}^{\mathrm{Ser}}_{\mathrm{UCA}}$, and also the presumably similar or identical products of the less well-characterized $\mathrm{tRNA}^{\mathrm{Ser}}_{\mathrm{minor}}$ genes on *Eco*RI fragments *A* and *D*) is intriguing. The original wobble hypothesis allowed U·G pairings between a U in the first position of the anticodon and a G in the last position of the codon[36]. However, unmodified Us are almost unknown in this position of RNAs and there is evidence that one of the effects of the ubiquitous modification of first position uridines is to restrict translational specificity[37]. In the case of the yeast $\mathrm{tRNA}^{\mathrm{Ser}}_{\mathrm{minor}}$ species with a U*GA anticodon, the U* is probably a derivative of 5-carboxymethyluridine as this is found in alkaline hydrolysates of $\mathrm{tRNA}^{\mathrm{Ser}}_{\mathrm{minor}}$ (P.W.P., unpublished results). Derivatives of 5-carboxymethyluridine have been found in the first positions of the anticodons for yeast $\mathrm{tRNA}^{\mathrm{Lys}}_{2}$, $\mathrm{tRNA}^{\mathrm{Arg}}_{3}$ and $\mathrm{tRNA}^{\mathrm{Glu}}_{3}$, and in all three cases triplet binding studies indicate a strong preference for A over G in the last position of the corresponding codons[38–40]. In view of these precedents, the hypothesis that the U*GA anticodon is largely restricted to the translation of UCA is reasonable. This restriction would explain why *sup-RL1*$^+$ is an essential gene because its gene product would presumably be the only $\mathrm{tRNA}^{\mathrm{Ser}}$ that could efficiently translate UCG. The evolution of a special class of UCA-specific tRNAs suggests that the IGA anticodon of the major $\mathrm{tRNA}^{\mathrm{Ser}}_{2}$ species may be poorly suited to the translation of UCA. A hint that UCA codons lead to suboptimal translational efficiency in yeast is found in the DNA sequences of two genes that code for the highly abundant glycolytic enzyme glyceraldehyde-3-phosphate dehydrogenase: in 50 occurrences of serine codons, only UCU and UCC are used[41].

Whatever the selective pressures that gave rise to this gene family, the close sequence relationship (96% homology) between $\mathrm{tRNA}^{\mathrm{Ser}}_{\mathrm{UCG}}$ and $\mathrm{tRNA}^{\mathrm{Ser}}_{\mathrm{UCA}}$ reflects an unusual evolutionary phenomenon: these genes have diverged functionally with the alteration of only three base pairs. Two of these alterations seem to be functionally neutral as their only consequence is a change from a G·C to C·G base pair in the anticodon stems of the tRNA products. The third alteration, however, affects the functionally critical site that codes for the wobble position base of the tRNA's anticodon. The presence of an intervening sequence in one of these genes (*SUP-RL1*) and its absence in the other (*SUQ5*) is intriguing. Like the two non-allelic rat insulin genes[42], the yeast $\mathrm{tRNA}^{\mathrm{Ser}}_{\mathrm{UCG}}$ and $\mathrm{tRNA}^{\mathrm{Ser}}_{\mathrm{UCA}}$ genes represent a situation in which one gene in a closely related pair has either gained or lost an intervening sequence during an evolutionary period that was too short to allow substantial coding region divergence. These cases, as well as the existence of a strain-dependent pattern of intervening sequence occurrence in rRNA genes in the *Tetrahymena* micronucleus[43,44] and yeast mitochondria[45,46] cast doubt on the hypothesis that only extremely rare evolutionary events can result in the acquisition or loss of an intervening sequence from a functional eukaryotic gene[47,48].

Note added in proof: Two recent studies support the view of the $\mathrm{tRNA}^{\mathrm{Ser}}$ gene family described here, and they provide important new details about the $\mathrm{tRNA}^{\mathrm{Ser}}_{\mathrm{UCA}}$ genes: a third suppressor of the *SUQ5* type, *SUP17*, has been discovered[51] and the three suppressors *SUQ5*(=*SUP16*), *SUP17* and *SUP19* have been individually correlated with a set of three $\mathrm{tRNA}^{\mathrm{Ser}}_{\mathrm{minor}}$-hybridizing *Hin*dIII fragments, which are the counterparts of our *Eco*RI fragments *A*, *B* and *D* (J. R. Broach, L. R. Friedman and F. Sherman, personal communication).

Received 23 October 1980; accepted 2 April 1981.

1. Hawthorne, D. C. & Mortimer, R. K. *Genetics* **48,** 617–620 (1963).
2. Stewart, J. W. & Sherman, F. *J. molec. Biol.* **68,** 429–443 (1972).
3. Stewart, J. W., Sherman, F., Jackson, M., Thomas, F. L. X. & Shipman, N. *J. molec. Biol.* **68,** 83–96 (1972).
4. Capecchi, M. R., Hughes, S. H. & Wahl, G. M. *Cell* **6,** 269–277 (1975).
5. Gesteland, R. F. *et al. Cell* **7,** 381–390 (1976).
6. Piper, P. W. *et al. Nature* **262,** 757–761 (1976).
7. Goodman, H. M., Olson, M. V. & Hall, B. D. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5453–5457 (1977).
8. Hopper, A. K., Banks, F. & Evangelidis, V. *Cell* **14,** 211–219 (1978).
9. Knapp, G., Beckmann, J. S., Johnson, P. F., Fuhrman, S. A. & Abelson, J. *Cell* **14,** 221–236 (1978).
10. O'Farrell, P. Z., Cordell, B., Valenzuela, P., Rutter, W. J. & Goodman, H. M. *Nature* **274,** 438–445 (1978).
11. Olson, M. V., Hall, B. D., Cameron, J. R. & Davis, R. W. *J. molec. Biol.* **127,** 285–295 (1979).
12. Kurjan, J., Hall, B. D., Gillam, S. & Smith, M. *Cell* **20,** 701–709 (1980).
13. Olson, M. V., Loughney, K. & Hall, B. D. *J. molec. Biol.* **132,** 387–410 (1979).
14. Zachau, H. G., Dütting, D. & Feldmann, H. *Z. phys. Chem.* **347,** 212–235 (1966).
15. Piper, P. W. *J. molec. Biol.* **122,** 217–235 (1978).
16. Etcheverry, T., Colby, D. & Guthrie, C. *Cell* **18,** 11–26 (1979).
17. Kruppa, J. & Zachau, H. G. *Biochim. biophys. Acta* **277,** 499–512 (1972).

18. Brandriss, M. C., Stewart, J. W., Sherman, F. & Botstein, D. *J. molec. Biol.* **102,** 467–476 (1976).
19. Ono, B.-I., Stewart, J. W. & Sherman, F. *J. molec. Biol.* **128,** 81–100 (1979).
20. Cox, B. S. *Heredity* **20,** 505–521 (1965).
21. Cox, B. S. *Heredity* **26,** 211–213 (1971).
22. Liebman, S. W., Stewart, J. W. & Sherman, F. *J. molec. Biol.* **94,** 595–610 (1975).
23. Piper, P. W. & Wasserstein, M. *Eur. J. Biochem.* **80**, 103–109 (1977).
24. Murray, N. E., Brammar, W. J. & Murray, K. *Molec. gen. Genet.* **150,** 53–61 (1977).
25. Hicks, J. & Fink, G. R. *Nature* **269,** 265–267 (1977).
26. Smith, H. O. & Birnstiel, M. L. *Nucleic Acids Res.* **3,** 2387–2398 (1976).
27. Clarkson, S. G., Kurer, V. & Smith, H. O. *Cell* **14,** 713–724 (1978).
28. Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74,** 560–564 (1977).
29. Valenzuela, P., Venegas, A., Weinberg, F., Bishop, R. & Rutter, W. J. *Proc. natn. Acad. Sci. U.S.A.* **75,** 190–194 (1978).
30. Knapp, G., Ogden, R. C., Peebles, C. L. & Abelson, J. *Cell* **18,** 37–45 (1979).
31. Ogden, R. C. *et al. Cell* **17,** 399–406 (1979).
32. Venegas, A., Quiroga, M., Zaldivar, J., Rutter, W. J. & Valenzuela, P. *J. biol. Chem.* **254,** 12306–12309 (1979).
33. Müller, F. & Clarkson, S. G. *Cell* **19,** 345–353 (1980).
34. Page, G. S. thesis, Univ. Washington (1979).
35. Ginsberg, T., Rogg, H. & Staehelin, M. *Eur. J. Biochem.* **21,** 249–257 (1971).
36. Crick, F. H. C. *J. molec. Biol.* **19,** 548–555 (1955).
37. McCloskey, J. A. & Nishimura, S. *Acc. chem. Ref.* **10,** 403–410 (1977).
38. Kobayashi, T., Irie, T., Yoshida, M., Takeishi, K. & Ukita, T. *Biochim. biophys. Acta* **366,** 168–181 (1974).
39. Sen, G. C. & Ghosh, H. P. *Nucleic Acids Res.* **3,** 523–535 (1976).
40. Weissenbach, J. & Dirheimer, G. *Biochim. biophys. Acta* **518,** 530–534 (1978).
41. Holland, J. P. & Holland, M. J. *J. biol. Chem.* **255,** 2596–2605 (1980).
42. Lomedico, P. *et al. Cell* **18,** 545–558 (1979).
43. Din, N., Engberg, J., Kaffenberger, W. & Eckert, W. A. *Cell* **18,** 525–532 (1979).
44. Wild, M. A. & Gall, J. G. *Cell* **18,** 533–543 (1979).
45. Bos, J. L., Heyting, C., Borst, P., Arnberg, A. C. & VanBruggen, E. F. J. *Nature* **275,** 336–338 (1978).
46. Faye, G., Dennebouy, N., Kujawa, C. & Jacq, C. *Molec. gen. Genet.* **168,** 101–109 (1979).
47. Crick, F. *Science* **204,** 264–271 (1979).
48. Darnell, J. E. Jr *Science* **202,** 1257–1260 (1978).
49. Southern, E. M. *J. molec. Biol.* **98,** 503–517 (1975).
50. Brandriss, M. C., Soll, L. & Botstein, D. *Genetics* **79,** 551–560 (1975).
51. Ono, B.-I., Wills, N., Stewart, J. W., Gesteland, R. F. & Sherman, F. *J. molec. Biol.* (in the press).

LETTERS TO NATURE

Baryosynthesis and the origin of galaxies

John D. Barrow*† & Michael S. Turner†‡

* Department of Physics, University of California, Berkeley, California 94720, USA
† Institute for Theoretical Physics, University of California, Santa Barbara, California 93106, USA
‡ Astronomy and Astrophysics Center, The University of Chicago, Chicago, Illinois 60637, USA

In the gravitational instability theory, the present structure in the Universe is envisaged to have evolved from small initial inhomogeneities which are of two generic types: isothermal and adiabatic. It has been shown that in the grand unified theory (GUT) for baryosynthesis primordial isothermal perturbations are incompatible with the standard Friedman–Robertson–Walker (FRW) universe. Here we discuss four non-standard models in which the two are compatible. Of these models two prove to be viable: (1) a universe which is initially dominated by large-scale inhomogeneous shear and (2) a universe which undergoes a de Sitter phase during the GUT phase transition. We also show how to deduce new constraints on chaotic behaviour near the cosmological singularity.

Two outstanding fundamental problems in cosmology are the origin of matter and the origin of structure—stars, galaxies and clusters. Although the laws of physics are very nearly matter–antimatter symmetrical, the Universe seems to possess an overt asymmetry and contains a negligible amount of antimatter[1]. The average ratio of baryons to 3K microwave photons[2] is $10^{-10\pm1}$; the baryon-to-specific entropy ratio (kn_{B}/s), which remains constant when baryon number n_{B} is conserved and the expansion is isentropic, combines these two observations[2], $B \equiv kn_{\mathrm{B}}/s = 10^{-11\pm1}$. In the gravitational instability picture (reviewed in refs 3, 36), the present large-scale structure developed from small-density inhomogeneities which existed earlier but the origin of these fluctuations is a key problem. These perturbations are of two generic types: adiabatic, in which both the matter and radiation participate and which are isentropic, $\delta B = 0$, and isothermal, which during the radiation-dominated epoch are simply spatial fluctuations in B. During the radiation epoch adiabatic perturbations on scales larger than the horizon ($|\mathbf{x}| \geqslant t$) are gravitationally unstable and amplify[3,36] in co-moving proper time, $\delta_{\mathrm{A}} \equiv \delta\rho/\rho \propto t$, while isothermal perturbations, $\delta_{\mathrm{I}} \equiv \delta \mathrm{B}/\mathrm{B}$, remain constant. In each case, when the matter and radiation decouple ($t \sim 10^{13}$ s) density perturbations on scales greater than the Jeans mass at decoupling are Jeans unstable and grow to produce the present structure. (The possibility of massive neutrinos modifies this picture, see ref. 4; but for present purposes these changes are not important.) There is no compelling evidence for the initial perturbations to be of any one type. However, if the initial perturbations were adiabatic, then variations in the microwave background[5] of $\mathrm{O}(10^{-4})$ should be observed on scales $\geqslant \mathrm{O}(30'')$. The present upper limits on such variations are very close to the expected variations[5]. The details of galaxy formation depend critically on which of these types of primordial perturbations were present.

The solution to one of these fundamental problems seems to preclude, or at least exacerbate, the solution to the other. For example, first, theories[6,7] in which the microwave background is of recent origin (cold universe models) have been advocated to explain the origin of structure; however, they offer no clue as to why the Universe contains only matter whilst the laws of physics are nearly matter–antimatter asymmetrical. Second, great progress has been made recently in understanding the origin of $kn_{\mathrm{B}}/s \simeq 10^{-11}$ within the context of the standard hot big-bang model[8]. It has been suggested[10–15] that the out-of-equilibrium decay of super massive ($\geqslant 10^{14}$ GeV) gauge or Higgs bosons which are predicted by GUTs, and whose interactions violate B, C and CP, can lead to the production of a baryon excess of the required magnitude ($B \simeq 10^{-11}$). Detailed numerical calculations[16,17] show that this type of explanation is predictive and can work. However, in the standard FRW model of the hot universe the GUTs preclude the production of isothermal perturbations during baryosynthesis—which could be a serious limitation in attempts to explain the formation of large-scale structure in the universe[18].

We now examine four non-standard models in which it is possible to have both baryosynthesis and isothermal perturbations. The first two have been suggested elsewhere and involve curvature fluctuations[7] and primordial black holes[18–20]. In the third the dynamics of the universe are initially dominated by large-scale inhomogeneous shear, and in the fourth the universe undergoes a de Sitter phase due to a first-order phase transition[21–24].

We first need to explain why the GUTs scenario for baryosynthesis excludes the existence of isothermal perturbations in the standard hot big-bang model. Detailed calculations[16,17] show that the baryon asymmetry produced by superheavy boson decay is given by

$$B = 10^{-2}\varepsilon f(\alpha, K) \tag{1}$$

where ε measures the CP violation in the superheavy system ($\equiv$ mean net baryon number produced by the decay of a super heavy boson–antiboson pair), α is the coupling strength of the superheavy, and $K = \Gamma_{\mathrm{D}}/H$ for $T = M$(Γ_{D} is the decay rate of the superheavy boson, M is the mass of the superheavy boson, and $H \equiv \dot{R}/R$ is the expansion rate of the universe) is a measure of the deviation from local thermal equilibrium. In the FRW universe[25] $H = (8\pi G\rho/3)^{1/2}$ and $\rho = g_*(T)(\pi^2/30)T^4$, so that

$$H^2 = \frac{8\pi G\rho}{3} g_* \frac{\pi^2}{30} T^4 \equiv AT^4 \tag{2}$$

0028-0836/81/240469—04$01.00

where $g_*(T)$ counts the relativistic degrees of freedom ($g_* = \sum_{\text{bosons}} g + 7/8 \ \sum_{\text{fermions}} g$). $\Gamma_D = c\alpha M$, where $c \sim 0(1)$ and depends on the GUT. Therefore,

$$K = c\left(\frac{4\pi^3}{45}\right)^{-1/2} g_*(M)^{-1/2}\alpha(GM)^{-1/2} \tag{3}$$

so that B only depends on particle physics parameters and constants, and can have no spatial variation. In minimal SU(5) $K = 3\times10^{17}\alpha\,\text{GeV}/M$. For $K \ll 1$, $\Gamma_D \ll H$ when $T \simeq M$ and so superheavy bosons cannot diminish in number and remain overabundant (relative to their equilibrium abundance) for $T \leqslant M$. When they eventually do decay, ($T \ll M$), they produce a baryon excess. In this case, $f(\alpha, K) \simeq 1$ and the maximum asymmetry arises. For $K \gtrsim 1$, $\Gamma_D \gtrsim H$ and decays do occur when $T \simeq M$; an equilibrium abundance of superheavy bosons is maintained to some extent, and a lesser asymmetry evolves. For $1 \leqslant K \leqslant 30$, numerical results[17] reveal

$$f(\alpha, K) \simeq K^{-1.3} \tag{4}$$

(essentially independent of α). From equations (1), (3) and (4), to produce $\delta_I \neq 0$, one must either modify the time–temperature relationship (2) in a spatially dependent way, have a spatially dependent ε, or modify the GUTs scenario.

If the GUTs scenario for baryosynthesis is discarded, then arbitrary initial fluctuations in B can be specified. In the GUTs picture it is still possible to specify an initial $B(\mathbf{x})$; however, in most instances the initial $B(\mathbf{x})$ is destroyed by baryon nonconserving reactions[17]. If the initial $B(\mathbf{x})$ is associated with some quantity conserved by the GUT, then the initial fluctuations will survive—but this is not a very appealing possibility.

Neither the existence nor the survival of adiabatic inhomogeneities ($\delta_A \neq 0$) in the early universe depend on the GUTs. However, it might be possible to demonstrate that if adiabatic fluctuations are present in the early universe when baryosynthesis is occurring, then a spatial variation in K (~baryosynthesis efficiency) must occur. A large-scale ($|\mathbf{x}| > t$) adiabatic inhomogeneity is equivalent to a gradient in the spatial three-curvature of space-time ($\delta_A \sim \delta^{(3)}R$). If we suppose this curvature inhomogeneity to be spherically symmetric, then, in accord with Birkhoff's theorem[25], it will behave as an independent Friedmann universe embedded in a flat background. The temperature–time adiabat of the spatially flat background will be given by equation (2), and that of the inhomogeneity by

$$\tilde{H}^2 = A\tilde{T}^4 - C\tilde{T}^2 \tag{5}$$

where $\tilde{}$ denotes quantities parameterizing the inhomogeneity and C is the curvature parameter. The efficiency of baryosynthesis in the two regions can be computed from the parameters K and $\tilde{K}$:

$$\tilde{K}/K = \left(1 - \frac{C}{AM^2}\right)^{-1/2} \tag{6}$$

According to equation (4) we see that when the amplitude of the inhomogeneity is small and $K \gtrsim 1$, the ratio of specific entropies in the regions of differing curvature departs from unity

$$\tilde{B}/B \simeq 1 - 0.65\frac{C}{AM^2} \tag{7}$$

In general $C(\mathbf{x})$ is some fairly arbitrary function of space (or mass scale). Where $C > 0$, the overdense regions expand more slowly than the flat ($C = 0$) background, equilibrium is more easily maintained, and consequently a smaller baryon number is generated than in the background. Note that this inhomogeneous distribution of baryons occurs when $K \gtrsim 1$, $M \leqslant 10^{16}$ GeV for gauge bosons and $M \leqslant 3\times10^{13}$ GeV for Higgs bosons, and also that regions of radiation excess ($C > 0$) finally emerge with a baryon deficit. Note that $f(\alpha, K)$ has a slight K-dependence for $K < 1$, so that small variations in B will arise even if $K < 1$.

Although this argument indicates that, in principle, adiabatic modes generate isothermal ones of similar amplitude and spectrum at $t_{\text{GUT}} \sim 10^{-35}$ s, in practice these isothermal modes will be physically irrelevant to the origin of galaxies and clusters. During the subsequent period of universal expansion between $t \sim 10^{-35}$ s and the end of the radiation-dominated era, $t \sim 10^{12}$ s, the isothermal modes will not grow, whilst the accompanying adiabatic modes will grow, $\delta_A \propto t$, until they enter the horizon. By the epoch of galaxy formation the amplitude of any isothermal modes will be utterly negligible compared with any adiabatic ones that were of similar amplitude at t_{GUT}.

If curvature fluctuations have large amplitude ($C/AM^2 \sim 1$) then equation (5) indicates that $\tilde{H}$ may fall to zero. If this occurs when $|\mathbf{x}| \sim t$ then black hole formation is inevitable. It has been suggested that this creates another means of producing isothermal fluctuations[18,19]:

Black holes smaller than $m_* \sim 4\times10^{-2}\, m_p^2 M^{-1}$ will have a Bekenstein–Hawking temperature, T_{BH}, in excess of M($m_p = G^{-1/2} \simeq 1.2\times10^{19}$ GeV). When they evaporate, a significant portion of their mass is radiated as superheavy boson ($X\bar{X}$) pairs. The $X\bar{X}$ ejecta decay out of equilibrium and generate an inhomogeneous distribution of baryons through their baryon-asymmetric decays. The amplitude and spectrum of the final baryon inhomogeneity are simply related to those of the adiabatic perturbations that collapsed into black holes. Again, an isothermal spectrum can be generated but the subsequent growth of the adiabatic perturbations associated with the formation and explosion of the holes will render the isothermal modes completely insignificant in the late stages of expansion.

Another type of inhomogeneity could be present in the earliest stages of the universe ($t \leqslant 10^{-35}$ s); one which does not suffer from the above-mentioned difficulties, and which, in the course of expansion will decay away adiabatically as t^{-1}. Suppose the early universe were to contain an inhomogeneous distribution of shear (gravitational wave) anisotropy, that is, at every point $\mathbf{x}$ the three orthogonal expansion rates would differ when measured at the same proper time from the initial singularity. When the shear inhomogeneity is large scale ($|\mathbf{x}| > t$), then a good description of the behaviour of the whole space–time is given by the velocity-dominated approximation[26–28]. If we assume the singularity is simultaneous in space at $t = 0$, then the temperature–time relationship is a weighted combination of Kasner and Friedmann behaviours,

$$\rho = (g_*\pi^2/30)T^4 = \frac{3}{32\pi G}\, t^{-4/3}(t + t_*(\mathbf{x}))^{-2/3} \tag{8}$$

where the arbitrary spatial function $t_*(\mathbf{x})$ indicates the time when the expansion at $\mathbf{x}$ passes from being anisotropic to roughly isotropic (t_* is directly proportional to the gravitational wave energy at $\mathbf{x}$). Equation (8) describes a universe that begins ($t \ll \bar{t}_*$) as an inhomogeneous Kasner universe but expands to approach a flat Friedmann model to a better and better approximation in the course of time, $t \gg \bar{t}_*$ (overbar here denotes the mean value of the spatial function $t_*(\mathbf{x})$.) It is known that exact solutions to the Einstein equations exist which possess inhomogeneous shear anisotropy of this form. However, it is not known whether this behaviour is generic to cosmological solutions of the Einstein equations or whether curvature fluctuations must generically co-exist with shear fluctuations.

Suppose the shear inhomogeneity is large at the time of baryosynthesis, $t_{\text{GUT}} \ll \bar{t}_*$, then the shear distribution creates a spatial variation in K. We find

$$K(\mathbf{x}) \propto t_*^{-1/2}(\mathbf{x}) \tag{9}$$

If K exceeds unity, baryosynthesis will reflect this inhomogeneity, and create isothermal fluctuations with

$$\delta_I = (2/3)\delta M^{-\alpha} \tag{10}$$

where $t_*(\mathbf{x}) = \bar{t}_*(1 + \delta M^{-\alpha})$ describes the spectrum of inhomogeneous shear on the mass (length) scale $M(|\mathbf{x}|)$. Although the dependence $B \propto K^{-1.3}$ strictly only applies in the isotropic case, we have assumed that it is also a reasonable description in the anisotropic case. When the amplitude of the shear inhomogeneity is small, $t_{\text{GUT}} \gg \bar{t}_*$ a similar trend is obtained

$$\begin{aligned} K(\mathbf{x}) &\propto 1 - \frac{1}{9}[t_*(\mathbf{x})/t_{\rm GUT}]^2 \\ \delta_{\rm I} &\simeq 0.3\delta(\bar{t}_*/t_{\rm GUT})^2 M^{-\alpha} \end{aligned} \qquad (11)$$

(if $K < 1$ lesser fluctuations arise).

The shear inhomogeneity will subsequently decay away in accord with equation (8), but the baryon fluctuations will remain to constitute an isothermal spectrum of density inhomogeneities in the later stages of the expansion. It is possible to 'explain' any desired isothermal spectrum by a suitable choice of the gravitational wave energy density $t_*(\mathbf{x})$—a situation no less arbitrary than in all other models where the isothermal spectrum is given directly by initial conditions. (Note that a spectrum of primordial rotational inhomogeneities could generate baryon inhomogeneity but the rotational perturbations will subsequently grow in amplitude relative to the radiation density as $\omega^2/\rho \propto t$ until the end of the radiation era; ω is the angular velocity. If they were large enough to generate significant primordial baryon number fluctuations they would now be too large to be compatible with present microwave background isotropy measurements.)

There is another interesting consequence of this generation process. Suppose the very early ($t \leqslant 10^{-35}$ s) universe were chaotic in structure[29,30], dominated by inhomogeneous and anisotropic motions as described by a relation of the form of equation (8) where $t_*(\mathbf{x}) \gg t$. In the absence of the GUTs scenario, the chaos would decay away $\propto t^{-1}$ in the course of time and this information concerning the very earliest stages of the universe would be unavailable. However, if B evolves through the GUTs picture, then the chaotic dynamics before time t_* are imprinted in the baryon distribution and this information is not redshifted away. In particular, shear fluctuations of large amplitude ($\nabla_x t_* \gg 1$) on a scale $\mathbf{x}$ would produce baryon inhomogeneities of such large contrast ($\delta_{\rm I} \geqslant 1$) that at the end of the radiation era, when freed from the radiation background, they would form a profusion of supermassive black holes on all scales $\mathbf{x}$ exceeding the Jeans mass at $t_{\rm rec} \sim 10^5$–$10^6 M_\odot$. The strong observational limits[31,32] on the presence of large black holes indicate that the universe was not dominated by very large-amplitude gravitational wave inhomogeneities at $t \sim 10^{-35}$ s. If small-amplitude isothermal modes were to arise from shear inhomogeneities, then the geometry of the baryon distribution on very large scales (~100 Mpc) might be anticipated to reflect the ellipsoidal configurations of the gravitational wave modes. Pronounced flattening of very large-scale structures might be expected which would reflect the dynamics of the universe at $t_{\rm GUT}$ rather than local gravitational instabilities which only generate deviations from spherical symmetry in much smaller structures that undergo collapse.

Finally, there is another way to produce primordial isothermal fluctuations within the GUTs scenario. From equation (1), B depends only on α, K and ε. We have already discussed the consequences of producing spatial variations in K. We will now consider how spatial variations in ε might give rise to isothermal perturbations. The CP violation ε is one of two generic types: (1) hard, in which CP-violating terms appear explicitly in the lagrangian, and (2) soft, where the violation is a consequence of spontaneous symmetry breaking. In the former, ε is constant throughout the universe, and so could not give rise to $\delta_{\rm I} \neq 0$. In the latter, ε depends on phase angles 'chosen' by the symmetry breaking. These should be uncorrelated over dimensions larger than the particle horizon at the time of symmetry breaking ($t \sim t_{\rm GUT}$). (The analogue in ferromagnetism is the different directions of magnetization in independent domains of a ferromagnet.)

In most models incorporating soft CP violation, the sign of ε varies from domain to domain, so that the universe is on the whole matter–antimatter symmetrical (up to purely statistical fluctuations)[33]. Yanagida and Yoshimura[34] have suggested that only the magnitude (and not the sign) of ε may vary. We might call this 'semi-soft' CP violation. (One can create an *ad hoc* model by considering the possibility of having both hard and soft CP violation, so that $\varepsilon = \varepsilon_{\rm H} + \varepsilon_{\rm s}$.) In this case we would have

$$\varepsilon(\mathbf{x}) = \varepsilon_0[1 + \delta \operatorname{sgn}(\cos\theta(\mathbf{x}))] \qquad (12)$$

where $\theta(\mathbf{x})$ varies randomly on scales larger than a domain. So we expect baryosynthesis to create a Poisson spectrum of isothermal perturbations,

$$\delta_{\rm I} \simeq (M/M_{\rm D})^{-1/2} \qquad (13)$$

where $M_{\rm D}$ is the baryon mass contained in a domain at $t_{\rm GUT}$. The difficulty is that a domain can be no larger than a region which can internally communicate. In the FRW model this is just the volume within the horizon $\simeq (4\pi/3) \times (2t)^3$, so that the number of baryons contained in this volume is

$$\begin{aligned} N_{\rm B} &= (kn_{\rm B}/s)(s/k)(32\pi/3)t^3 \\ &= 10^{-11\pm1}(5/\pi^3)^{1/2} g_*^{-1/2}(T)(m_{\rm p}/T)^3 \end{aligned} \qquad (14)$$

For $T \simeq 10^{15}$ GeV and $g_* \sim 100$, $N_{\rm B} \simeq 10^{\pm1} \simeq 10^{-57} M_\odot$. The isothermal fluctuations on the scales of interest ($M \geqslant 10^5 M_\odot$) are $\leqslant 0(10^{-31})$ which is too small to be of any significance.

However, it has been suggested[21–24] that the GUT phase transition may be first order, and that if the transition is slow enough, then the universe goes through a de Sitter phase. During this period the horizon grows exponentially, $d_{\rm H} \propto \exp(t/\tau)$. If the transition takes $\sim 100 t_{\rm GUT}$, then $M_{\rm D}$ could easily exceed $10^5 M_\odot$. If this were the case, the semi-soft scenario could provide the required spectrum of isothermal perturbations. (Numerical simulations have demonstrated that $\delta_{\rm I} \simeq (M/10^6 M_\odot)^{1/2}$ can reproduce the present structure[3,36].) There is no way at present to calculate whether or not the de Sitter phase should occur or how long its duration would be if it did. In addition there seem to be difficulties associated with the transition back to a FRW universe from the de Sitter phase[21], although Eardley and Press (personal communication) may have found a solution to this difficulty.

Although in the standard FRW model baryosynthesis and isothermal protogalaxies are mutually incompatible, we have suggested two scenarios in which they are both allowed. In the first, the universe must initially be dominated by large-scale inhomogeneous shear, and the superheavy boson responsible for baryosynthesis should be less massive than $\simeq 3 \times 10^{17} \alpha$ GeV. The spectrum of isothermal perturbations mirrors that of the inhomogeneous shear. In the second, the CP violation ε must be semi-soft and the universe must undergo a de Sitter phase. A Poisson spectrum of isothermal fluctuations is produced.

We thank R. Bond, D. Eardley, C. J. Hogan, E. Kolb and J. Silk for helpful discussions. This work was supported in part by the Institute for Theoretical Physics (NSF grant PHY77-27084), The Miller Institute, DOE grant AC02-80ER10773 and an SRC grant to M.S.T. from the Department of Astrophysics, Oxford.

Received 30 January; accepted 24 April 1981.

1. Steigman, G. *A. Rev. Astr. Astrophys.* **14,** 339 (1976).
2. Olive, K. A., Schramm, D. N., Steigman, G., Turner, M. S. & Yang, J. *Astrophys. J.* **246** (1980).
3. Barrow, J. D. *Phil. Trans. R. Soc.* A**296,** 273 (1980).
4. Bond, R., Efstathiou, G. & Silk, J. *Phys. Rev. Lett.* **45,** 1980 (1980).
5. Wilson, M. L. & Silk, J. *Astrophys. J.* **243,** 14 (1981).
6. Rees, M. J. *Nature* **275,** 35 (1978); *Phys. Rev. Lett.* **28,** 1669 (1972).
7. Carr, B. J. & Rees, M. J. *Astr. Astrophys.* **61,** 705 (1977).
8. Barrow, J. D. *Surv. High Energy Phys.* **1,** 183–212 (1980).
9. Sakharov, A. D. *JETP Soviet Phys. Lett.* **5,** 24 (1967).
10. Yoshimura, M. *Phys. Rev. Lett.* **41,** 281 (1978).
11. Ignatiev, A., Krasnikov, N., Kuzmin, V. & Taukhelidze, A. *Phys. Lett.* **76B,** 436 (1978).
12. Dimopoulos, S. & Susskind, L. *Phys. Rev.* D**18,** 4500 (1979).
13. Toussaint, D., Treiman, S. B., Wilczek, F. & Zee, A. *Phys. Rev.* D**19,** 1036 (1979).
14. Weinberg, S. *Phys. Rev. Lett.* **42,** 850 (1979).
15. Ellis, J., Gaillard, M. K. & Nanopoulos, D. V. *Phys. Lett.* **80**B, 360 (1978).
16. Kolb, E. & Wolfram, S. *Phys. Lett.* **91B,** 217 (1980); *Nucl. Phys.* **B172,** 224 (1980).
17. Fry, J. N., Olive, K. A. & Turner, M. S. *Phys. Rev. Lett.* **45,** 2074 (1980); *Phys. Rev.* D**22,** 2953 (1980); 2977 (1980).
18. Turner, M. S. & Schramm, D. N. *Nature* **279,** 303 (1979).
19. Barrow, J. D. *Mon. Not. R. astr. Soc.* **192,** 427 (1980).
20. Turner, M. S. *Phys. Lett.* **89B,** 155 (1980).
21. Guth, A. *Phys. Rev.* D**23,** 347 (1981).
22. Kazanas, D. *Astrophys. J. Lett.* **241,** L59–L63 (1980).
23. Sato, K. *Mon. Not. R. Astr. Soc.* (in the press); *Nordita preprint* 80/31 (1980).
24. Linde, A. *Rep. Prog. Phys.* **42,** 389 (1979).

25. Weinberg, S. *Gravitation and Cosmology* (Wiley, New York, 1972).
26. Tomita, K. *Prog. theor. Phys.* **48**, 1503 (1972).
27. Eardley, D. M., Sachs, R. K. & Liang, E. T. P. *J. math. Phys.* **13**, 99 (1972).
28. Barrow, J. D. *Mon. Not. R. astr. Soc.* **178**, 625 (1977).
29. Misner, C. W. *Astrophys. J.* **151**, 431 (1968).
30. Barrow, J. D. *Nature* **267**, 117 (1976).
31. Carr, B. J. *Commt. Astrophys.* **5**, 161 (1978).
32. Barrow, J. D. & Silk, J. *Gen. Rel. Grav.* **10**, 633 (1979).
33. Senjanovic, G. & Stecker, F. W. *Phys. Lett.* **96B**, 285 (1980).
34. Yanagida, T. & Yoshimura, M. Preprint 80/211 (Tohoku University, 1980).
35. Hogan, C. *Mon. Not. R. astr. Soc.* **185**, 889 (1978).
36. Peebles, P. J. E. *The Large Scale Structure of the Universe* (Princeton University Press, 1980).

Impact fusion and the field emission projectile

E. R. Harrison

Carter Observatory, PO Box 2909, Wellington, New Zealand

A conceivable approach to the problem of achieving controlled thermonuclear power is by impact fusion[1–3]. Small projectiles (macrons[4]), of mass ~0.1 g and of appropriate material and design, are accelerated to a velocity $v \sim 10^8$ cm s^{-1}; they then collide with a deuterium–tritium target, and their kinetic energy is abruptly converted into thermal energy (~10 keV per nucleon, temperature ~10^8 K) that is inertially confined in a shocked region. Preliminary studies[1,2] in the early 1960s showed that in these conditions the released nuclear energy exceeds the initial kinetic energy. Despite its attractiveness, relatively little research has been devoted to the development of impact fusion as a power-producing technology, and only recently[5] have serious attempts been made to evaluate its importance compared with the more familiar proposed methods of magnetic and inertial confinement. The principal block to impact fusion is the development of accelerators capable of accelerating macrons to velocities two orders of magnitude greater than can be attained by existing methods. It is shown here, using general arguments, that the minimum length of such macron accelerators must exceed 1 km. A possible way of achieving the required high velocity is to accelerate, in a travelling electric field, macrons that are self-charged by electron field emission.

With a uniform acceleration $\dot{v}$, a nonrelativistic accelerator has a length

$$x = v^2/2\dot{v} \tag{1}$$

The differential or tidal stresses caused by an accelerating force F acting on a projectile of characteristic size L, are $\sim F/L^2$. If these stresses are not to exceed the yield stress s of the projectile, then the minimum length of the accelerator is

$$x_{\min} \sim \rho L v^2/s \tag{2}$$

for a projectile of density ρ and mass $\sim \rho L^3$. The criterion[1] for a net energy gain in impact fusion is approximately $\rho L > 0.1$ g cm^{-3}. With a reasonable value for the yield stress, of $s \sim 10^{10}$ dyn cm^{-2}, it is found that the minimum accelerator length is $x_{\min} \sim 1$ km. In practice, the actual minimum length might be ~ 10 km. This result is reasonably general and applies to electrostatic, magnetic (for example, rail guns[6]), travelling electromagnetic wave, beam-pushing[4], and laser ablation[4,7] accelerators. A few years ago an accelerator length of 10 km seemed prohibitive and impact fusion was hence deemed to be of academic interest only. But existing and proposed accelerators of subatomic particles are now on this scale[8], and because macron accelerators might conceivably be less costly, impact fusion has become a realistic possibility.

In spite of various problems (such as joule dissipation), it is widely thought that magnetic field acceleration is superior to electrostatic or travelling electric wave acceleration[7]. I shall attempt to show that electric field acceleration of self-charged macrons is capable of attaining the velocities required for impact fusion.

Consider first a conducting sphere of radius r, polarized in an electric field E, and having a net positive charge Q. The surface charge density is

$$\sigma = (Q + 3r^2E\cos\theta)/4\pi r^2 \tag{3}$$

where $\theta = 0$ lies in the direction of E. Imagine now a fine whisker of length h, small compared with r, protruding from the rear ($\theta = \pi$) of the sphere. The field strength at the tip of the whisker is

$$E_{\text{tip}} \sim -E(h/b)^2(1 - Q/Q_0) \tag{4}$$

where $Q_0 = 3r^2E$, and b is the radius of curvature at the tip of the whisker. With $E \sim 10^5$ V cm^{-1}, $h/b \sim 10^3$, the field emission current[9] is sufficient to create a charge $Q \sim Q_0$ on a time scale of 10^{-3} s. The acceleration is hence $\dot{v} = 9E^2/4\pi\rho r$, and the accelerator length is

$$x \sim \rho r v^2/E^2 \sim x_{\min} s/E^2 \tag{5}$$

where $x_{\min}$ is given by equation (2) with $L = r$. The maximum field intensity (at which disruption occurs) is $E_{\max} \sim s^{1/2}$, roughly equal to 3×10^7 V cm^{-1}. Usually, in practical applications we have $E \ll E_{\max}$, and therefore in this case the accelerator is exceedingly long.

Slightly better results are obtained when the trailing whisker has a length $h > r$. In this case the positive charge acquired by field emission is $Q_0 \sim rhE$, and the field strength necessary for maximum acceleration is hence $E_{\max} \sim (rs/h)^{1/2}$.

The above examples illustrate the proposed method, and show that the acceleration of spherical bodies charged by field emission is not very promising. The most striking results are obtained when we consider needle-shaped macrons. A prolate spheroidal body of $a \gg b$ (where a is the semi-major axis parallel to the applied electric field and b is the semi-minor axis), which is polarized and has a net positive charge Q, has an electric field at the trailing tip of

$$E_{\text{tip}} = (Q - Q_0)/b^2 \tag{6}$$

where

$$Q_0 = a^2E\Big/\left(\ln 2\frac{a}{b} - 1\right) \tag{7}$$

With a value of b sufficiently small, copious field emission of electrons occurs, and the positive charge Q tends to the value Q_0. The attained charge is about the same in value as that on a sphere of $r = a$, or

$$(Q_0/M)_{\text{ellipsoid}} \sim (a/b)^{4/3}(Q_0/M)_{\text{sphere}} \tag{8}$$

and the charge-to-mass ratio is increased by $(a/b)^{4/3}$ compared with a sphere of similar mass and density. The accelerator length in this case is

$$x = 2\pi b^2 \rho v^2\left(\ln 2\frac{a}{b} - 1\right)\Big/3aE^2 \tag{9}$$

The maximum possible applied electric field (at which disruption occurs) is found to be

$$E_{\max} \sim (b/a)s^{1/2} \tag{10}$$

thus yielding a minimum accelerator length in accord with equation (2) with $L \sim a$. Hence, when $a/b \sim 10^3$, only modest electric fields of $E \sim 10^4$ V cm^{-1} are required in a travelling wave linear accelerator of length 10 km. If the tensile strength of the whisker projectile could be as great as $s \sim 10^{12}$ dyn cm^{-2}, then stronger electric fields would offer the prospect of accelerators only hundreds of metres in length.

Field-emitting macrons have the following advantages: they are self-charging; large positive charges are in principle obtain-

---

* Permanent address: Department of Physics and Astronomy, University of Massachusetts, Amherst, Massachusetts 01003, USA.

0028–0836/81/240472—02$01.00

able; the net positive charge is concentrated towards the fore end and the macron is stably oriented; and the charge-to-mass ratio is predetermined thus allowing for synchronous acceleration in a travelling wave accelerator.

In general, macrons must fulfil two requirements: they must have a composition and configuration that allow optimum acceleration, and must have a configuration that yields a high temperature and density on impact with the target. Field-emitting macrons, which are needlelike in shape, do not necessarily have a configurational interaction with the target that allows for efficient use of their kinetic energy. For a needle-like macron the fusion criterion $\rho L \geqslant 0.1\ \mathrm{g\ cm^{-2}}$ applies with $b \leqslant L < a$, and a spherical macron having the same mass is undoubtedly superior[1]. We have therefore an apparent conflict between these two requirements. It may be possible, by means of an intense magnetic field of $B \sim s^{1/2} \sim 10^5$ G, to arrange for a change in configuration before impact with the target. Thus after acceleration, and before impact, the needle-like macron enters a region in which magnetic stresses cause it to collapse into a spherical body of radius $(ab^2)^{1/3}$.

The development of accelerators capable of accelerating macrons to high velocity will no doubt open up an entirely new field of 'macron physics'. Not only will it then be possible to study condensed matter at high temperatures, and perhaps achieve a controlled fusion technology, but with needle-like picogramme macrons it should also be possible to study relativistic effects ($v \sim c$) in macroscopic bodies.

I thank R. A. Alpher and E. Tadermaru for their comments, and M. N. Kreisler for advice and encouragement.

Received 6 February; accepted 16 April 1981.

1. Harrison, E. R. *Phys. Rev. Lett.* **11**, 1 (1963).
2. Winterberg, F. *Z. Naturforsch.* **19a**, 231 (1964).
3. Maisonnier, C. *Nuovo Cim.* **423**, 332 (1966).
4. Harrison, E. R. *Plasma Phys.* **9**, 183 (1967).
5. Peaslee, A. T. (ed.) *Proc. Impact Fusion Workshop*, LA-8000-C (Los Alamos Scientific Laboratory, 1979).
6. Marshall R. A. in *Proc. Impact Fusion Workshop*, LA-8000-C (Los Alamos Scientific Laboratory, 1979).
7. Ribe, F. L. & Vlases, G. C. in *Proc. Impact Fusion Workshop*, LA-8000-C (Los Alamos Scientific Laboratory, 1979).
8. Kreisler, M. N. in *Proc. Impact Fusion Workshop*, LA-8000-C (Los Alamos Scientific Laboratory, 1979).
9. Good, R. H. & Müller, E. W. *Handb. Phys.* **21**, 187 (1956).

British shore platforms and ice-sheets

J. B. Sissons

Department of Geography, University of Edinburgh, High School Yards, Edinburgh EH1 1NR, UK

During the past few years views about the relative ages and modes of formation of fossil rock platforms and cliffs in Scotland have changed considerably[1–3]. I suggest further changes here and present a model of rock-platform relationships which involves an interglacial set of platforms, a glacial set, and a late-glacial feature. Interpretations of the glaciation of the Outer Hebrides have recently been completely revised: this revision has major implications for the glaciation of the Inner Hebrides which, in turn, has wider implications.

It was accepted for a century that the last ice-sheet from the Scottish mainland crossed the Outer Hebrides, except for the mountain area of Harris, where it was confluent with local ice. It is now known that almost the whole of the Outer Hebrides was last covered by an independent ice-mass, which flowed towards the east and south-east over most of the island chain. The evidence for this flow comprises ice-moulded landforms, striae, and directions of transport of erratics from sources in the islands[4,5]. This means that the Scottish Highland ice-sheet terminated to the east of the Outer Hebrides. A stable ice margin in the mostly deep waters of the Minches and Sea of the Hebrides seems improbable. Relative stability of the Highland ice margin amidst the Inner Hebrides, where much of the terminal zone would have been land-based, seems more likely. One may envisage the ice advancing into the sea to the west from time to time owing to climatic deteriorations, these advances being followed by rapid retreat through calving during climatic ameliorations.

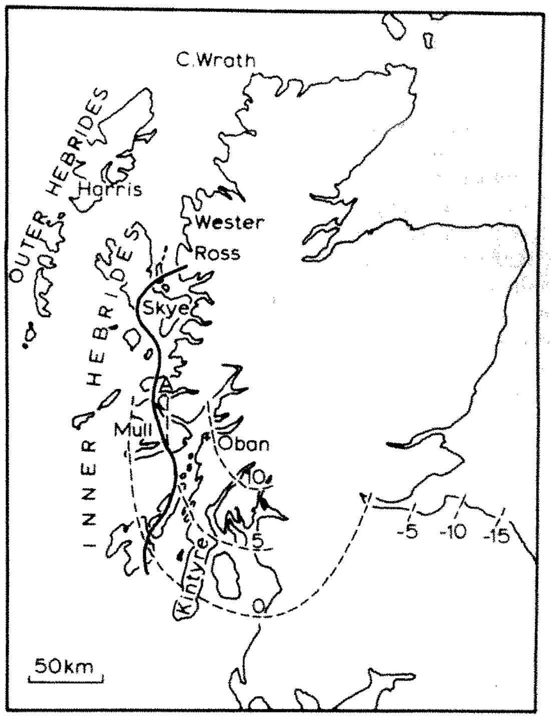


Fig. 1 Isobases for the Main Lateglacial Shoreline (after ref. 2) and eastern limit of high rock platforms as known at present (solid line). A, Ardnamurchan.

Growth of an ice-mass on the Outer Hebrides presents problems because erratics derived from sources outside the islands are present and because much of the area is rather low ground. I have therefore suggested[6] that the islands were covered by the last Highland ice-sheet (except for the mountains of Harris and surrounding ground). Subsequently, owing to calving, the major portion of the Highland ice-sheet was separated from ice on the Outer Hebrides. The situation described above should then have applied.

Shorelines cut in rock at reported altitudes of 18–51 m (mostly approximate and here converted to OD where necessary) are extremely well developed along many coasts in the Inner Hebrides. They also occur on the mainland coast in parts of Ardnamurchan and Wester Ross (Fig. 1)[7–9]. Some platform fragments are only 100–200 m wide but others are much broader, ranging up to 1 km. The backing cliffs are sometimes small, but others are tens of metres high, with a reported maximum of 90 m. Some platform fragments are striated and ice-moulded and some are overlain by till. This evidence of glaciation has led to the features being interpreted as formed during interglacials.

The interglacial interpretation encounters difficulties. (1) The required periods of prolonged interglacial sea-sevel stability are not recorded in the curves based on evidence from deep-sea cores[10]. (2) Because the platforms are much higher than platforms in England, Wales and Ireland that are regarded as interglacial, significant Quaternary uplift in the Inner Hebrides (unconnected with glacio-isostatic uplift) has been invoked. There is no independent evidence for such uplift. (3) 'High' platforms and cliffs are pronounced features in the Inner Hebrides and parts of the mainland west (and north) of the line shown in Fig. 1, but have not been reported east of the line.

Although glacial erosion can readily explain modification or removal of some platform fragments, it cannot alone account for this pronounced contrast.

In some limited areas high-platform fragments at quite different altitudes show clearly that more than one period of erosion is involved[8,11,12]. Despite this evidence the fragments have usually been discussed as if they constitute a single feature called 'the preglacial platform' or 'the interglacial platform'[7–9]. This has produced a simple error in reasoning: even though some platform fragments carry no evidence of glaciation it has led to the assumption that they have, nevertheless, been glaciated. Having thus assigned all rock-cut features to a 'preglacial' age, it has then been inferred that late-glacial coastal erosion 'scarcely affected solid rock' being 'restricted to the removal of older drift deposits' (ref. 11, p. 21). No convincing evidence of glaciation has been found on most platform fragments in Mull and northeastern Skye[8,13]. In Wester Ross some platform fragments form part of a late-glacial shoreline that elsewhere is represented by depositional forms (raised beaches and shingle ridges). Thus some platforms are not interglacial. Glaciated platforms also occur in Wester Ross but they are above the late-glacial marine limit[14,15].

It has been suggested that all the high-platform fragments in Wester Ross were formed by periglacial shore erosion[15]. It has also been argued that the 'high' platforms of western Scotland as a whole were produced during build-up of the last Scottish ice-sheet[3]. I now propose that the high-platform fragments represent a series of platforms produced during glacials. Frost riving would have been important, its effectiveness having been facilitated by removal of the resultant debris by waves and by repeated wetting in the intertidal zone during the period of each year when sea ice was absent. In sheltered areas (such as northeastern Skye and Wester Ross) frost riving would have been of major importance. In more exposed areas, where platforms are most extensive (for example, parts of western Mull), strong wave action would have operated in conjunction with frost action.

It is proposed that the altitudes of the features reflect glacio-isostatic uplift. The apparent absence of high platforms east of the line shown in Fig. 1 is attributed mainly to the presence of the Highland ice-sheet and of ice nourished on the high ground of eastern Mull and south-central Skye. The line (or one approximating to it) is interpreted as roughly representing the relatively stable position of the margin of this ice-mass, as mentioned above. I suggest that the platforms and cliffs were shaped and reshaped during successive glacials. Many glaciated platform fragments may be attributed to formation during ice-sheet advance[3] or readvance, when the relationships between eustatic sea-level and glacio-isostatic depression of the land differed from those during the last deglaciation of the localities in which they occur.

It has been inferred from deep-sea core evidence that major buildup of the North American and Scandinavian ice-sheets occurred at ~75 kyr ago[16]. If major buildup of the last Scottish ice-sheet also occurred at this time[3], there is likely to have been a long period during the last-glacial when 'high' platforms and cliffs could have been formed in western Scotland.

Along parts of the coastline of western Scotland there is a lower platform and cliff. The platform is typically 20–30 m wide, but sometimes exceeds 150 m, and the cliff is up to 15 m high. These features, often continuous for many kilometres, were initially interpreted as produced during the Holocene[8] but were subsequently regarded as interglacial[9,11,17,18]. I suggested the features were formed during the Loch Lomond Stadial (approximately equivalent to the Younger Dryas Stadial) and the later part of the preceding interstadial by frost action accompanied by limited wave action[1]. This view was doubted[19] but has been justified by detailed studies[20,21] which show that the shoreline (Main Lateglacial Shoreline) reaches ~11 m north-east of Oban, whence it slopes down towards west, southwest and south (Fig. 1). The feature has not been glaciated outside the limit of the Loch Lomond Advance but inside the limit only the cliff (or part of the cliff) is normally visible.

The Main Lateglacial Shoreline is also identified in south-eastern Scotland, where it is buried by later deposits or submerged by the sea. Its eastward slope has been traced from slightly above OD to −18 m (Fig. 1)[22–24]. Owing to less resistant rocks than in western Scotland, the rock platform is often hundreds of metres wide; in drift it may attain 2 km. Because late-glacial marine deposits are truncated by the plane of erosion and as deposits older than ~9,600 BP overlie it, a late-glacial age is strongly suggested. Such considerable planation requires a stable or more probably, a slowly transgressing sea. Stability followed by slow transgression during a period of 1,000–1,500 yr has been inferred, at the appropriate altitudes, from independent studies of marine fauna from sites in western Scotland[25,26].

Intertidal rock-platform fragments have been reported from many places in the Inner Hebrides and on the west and east coasts of the Scottish mainland. They are normally far less impressive features than those described above. Many of the fragments are ice-moulded or bear striae and some pass inland beneath cliffs of till. Interglacial formation is widely agreed. It has been recently proposed that intertidal platform fragments in part of the Inner Hebrides are parts of a single horizontal platform[2] but this view needs qualifying, first because intertidal platforms are obvious features owing to removal of any overlying deposits. They may therefore be a biased sample of a larger population. Second, where intertidal platforms, rising landwards, pass beneath till, the latter normally survives because it is above the reach of effective wave action. Because it is rarely possible to measure the inner edge of these higher, till-covered platforms the sample is again biased. Finally, in almost the whole area where detailed measurements have been made of the inner edge of intertidal fragments, the platform associated with the Main Lateglacial Shoreline also exists: in a considerable part of the area the latter is within a few metres of present sea-level. Consequently it is highly unlikely that interglacial platforms will have survived in the altitudinal zone affected by the later erosion. In this context it is very significant that in Kintyre, as the Main Lateglacial Shoreline descends to and passes below sea level, marine erosional features appear above it and become increasingly abundant southwards. They comprise raised cliffs, platforms, caves, stacks and geos, most of which are a few metres above high-water mark (~1.5 m OD). Evidence that the features have been glaciated is not firmly established, except for one rock platform (~13 m) that passes beneath a cliff of till[20,27–29].

The Main Lateglacial Shoreline has not been studied northwards from Mull and the Oban area but it may be expected to pass below present sea level in southern Skye. Significantly, in northern Skye there are rock platform fragments at ~2.5 and

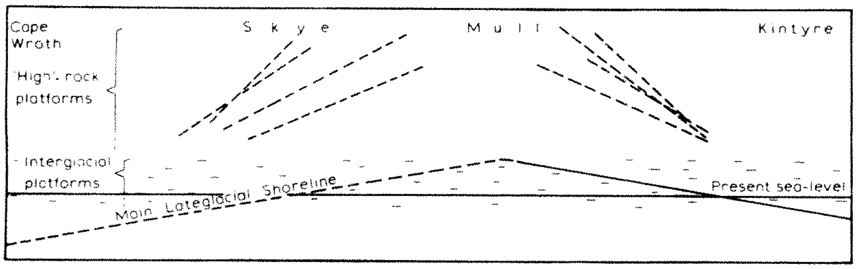


Fig. 2 Theoretical relationships of rock-cut shorelines along a north–south line in western Scotland.

~4 m while cemented beach deposits and the floors of cave entrances occur to altitudes several metres higher[13].

These relationships suggest that interglacial marine erosion features occur not only at present sea level but also for some distance above it. As areas immediately below low-water mark have not been investigated, such features may also continue below present sea level. Thus I envisage a zone of interglacial marine erosion. Individual features may have been formed during more than one interglacial, just as the glaciated intertidal platform is being reshaped during the present interglacial. This view accords with the deep-sea record, which shows that eustatic sea sevel attained different altitudes during different interglacials.

Figure 2 shows theoretical relationships in western Scotland for the three sets of evidence considered above. No vertical or horizontal scale is given as available altitudinal data on rock platforms are inadequate, except for the southward slope above sea level of the Main Lateglacial Shoreline. The number and relative gradients of the proposed high-rock platforms are diagrammatic. The high platforms should slope down to present sea level (and below it): evidence for this is lacking, probably because the Inner Hebrides extend insufficiently far westwards, while mainland areas where this might happen (for example, North-west Scotland) have not been studied in detail. In these latter areas it may prove difficult to distinguish high-rock platforms from interglacial ones. The interglacial platform fragments are shown in Fig. 2 as unconnected short lines as they are at present not correlated with each other: as a group the fragments suggest the existence of horizontal or gently inclined interglacial shorelines.

At many points around the coasts of South-west England, Wales and Ireland one or more shorelines cut in bedrock have been identified from slightly below sea-level to ~15 m altitude. In the glaciated area the rock platforms often pass beneath till and in unglaciated terrain beneath head. It is accepted that the platforms are interglacial but opinions differ as to the interglacial(s) involved[30–34]. The occurrence of interglacial platform fragments within a limited range of altitude in many parts of the British Isles suggests that they are one set of features: it is therefore suggested that the interpretation proposed for the Scottish features applies elsewhere. An implication of this interpretation is that glacio-isostatic recovery has essentially ceased (unless glacial erosion has yet to be compensated).

In Scottish areas so far studied the rock platform and cliff associated with the Main Lateglacial Shoreline are clear and distinctive. They are clearly preserved beneath present sea level in at least part of eastern Scotland and hence are likely to be identifiable below sea level elsewhere. I have suggested[35] that the Main Lateglacial Shoreline may be represented by the sharp break of slope traced for 30 km at −35 to −40 m in Cardigan Bay[36] and by the pronounced erosional shoreline at −42 m identified off part of the coast of South-west England[37]. The high rock platforms should relate to lower sea levels, perhaps of the order of −100 m.

If the interpretation of high rock platforms presented above is valid, then for much of the last glaciation (and, probably, other ones) the Scottish ice-sheet was far less extensive in western Scotland than previously envisaged. The ice that flowed east and south-east in the Outer Hebrides was at least 400 m thick at one point[38] ant it has left pronounced glacial erosion forms[5], implying that independent Outer Hebridean ice existed for a long time (or times). Many maps have been published showing an ice-sheet or ice-sheets terminating far to the south of Scotland (for example, at the latitude of south Wales or beyond). While such maps represent maximal conditions, they suggest that ice-sheets were normally extensive in the British Isles. The Hebridean evidence does not simply accord with this view for it suggests that, at least in Scotland, ice-sheets were normally of limited extent.

Received 20 February; accepted 23 March 1981.

1. Sissons, J. B. *Boreas* **3,** 41–48 (1974).
2. Dawson, A. G. *Proc. geol. Ass.* **91,** 339–344 (1980).
3. Sutherland, D. G. *Boreas* (in the press).
4. von Weymarn, J. thesis, Univ. Aberdeen (1974).
5. Flinn, D. *Geol. J.* **13,** 195–199 (1978).
6. Sissons, J. B. *Scott. J. Geol.* **16,** 81–84 (1980).
7. Wright, W. B. *Geol. Mag.* **48,** 97–109 (1911).
8. Bailey, E. B. *et al. Mem. geol. Surv.* (1924).
9. McCann, S. B. in *Geography at Aberystwyth* (eds Bowen, E. G., Carter, H. & Taylor, J. A.) 22–34 (University of Wales Press, Cardiff, 1968).
10. Shackleton, N. J. & Opdyke, N. D. *Quat. Res.* **3,** 39–55 (1973).
11. McCann, S. B. & Richards, A. *Scott. J. Geol.* **5,** 15–25 (1969).
12. Jardine, W. G. in *The Quaternary history of the Irish Sea* (eds Kidson, C. & Tooley, M. J.) 99–118 (Seel House Press, Liverpool, 1977).
13. Richards, A. *Scott. geogr. Mag.* **85,** 122–131 (1969).
14. Robinson, M. thesis, Univ. Edinburgh (1977).
15. Sissons, J. B. & Dawson, A. G. *Proc. Geol. Ass.* **92,** 115–124 (1981).
16. Ruddiman, W. F., McIntyre, A., Niebler-Hunt, V. & Durazzi, J. T. *Quat. Res.* **13,** 33–64 (1980).
17. McCallien, W. J. *Proc. Soc. Antiq. Scotl.* **71,** 174–206 (1937).
18. Gray, J. M. *Trans. Inst. Br. Geogr.* **61,** 81–99 (1974).
19. Peacock, J. D. *Scott. J. Geol.* **11,** 174–175 (1975).
20. Gray, J. M. *Trans. Inst. Br. Geogr.* **3,** 151–164 (1978).
21. Dawson, A. G. in *Studies in the Lateglacial of North-West Europe* (eds Lowe, J. J. Gray, J. M. & Robinson, J. E.) 45–53 (Pergamon, Oxford, 1980).
22. Sissons, J. B. *Trans. Inst. Br. Geogr.* **48,** 19–50 (1969).
23. Sissons, J. B. *Scott. geogr. Mag.* **92,** 17–29 (1976).
24. Eden, R. A., Carter, A. V. F. & McKeown, M. C. *Mar. Geol.* **7,** 235–251 (1969).
25. Peacock, J. D., Graham, D. K., Robinson, J. E. & Wilkinson, I. P. in *Studies in the Scottish Lateglacial environment* (eds Gray, J. M. & Lowe, J. J.) 89–100 (Pergamon, Oxford, 1977).
26. Peacock, J. D., Graham, D. K. & Wilkinson, I. P. *Rep. Inst. geol. Sci.* 78/17, 1978.
27. Nicol, J. *Q. Jl geol. Soc. Lond.* **8,** 406–425 (1852).
28. Hull, E. *Geol. Mag.* **3,** 5–10 (1866).
29. Sinclair, W. *Trans. geol. Soc. Glasg.* **14,** 170 (1913).
30. Wright, W. B. & Muff, H. B. *Scient. Proc. R. Dub. Soc.* **10,** 250–324 (1904).
31. George, T. N. *Proc. Geol. Ass.* **43,** 291–324 (1932).
32. Stephens, N. *Proc. R. Ir. Acad.* **58B,** 129–149 (1957).
33. Orme, A. R. *Field Stud.* **1,** 1–22 (1960).
34. Kidson, C. & Wood, R. *Proc. geol. Ass.* **85,** 223–237 (1974).
35. Sissons, J. B. *Nature* **280,** 199–203 (1979).
36. Dobson, N. C., Evans, W. E. & James, K. H. *Mar. Geol.* **11,** 27–69 (1971).
37. Kelland, N. C. *J. geol. Soc. Lond.* **131,** 7–17 (1975).
38. Peacock, J. D. *Scott. J. Geol.* **16,** 87–89 (1980).

Will the Sahelian drought end in 1985?

Hugues Faure

Laboratoire de Géologie du Quaternaire, CNRS, Faculté des
Sciences de Luminy Case 907, 13288 Marseille Cedex 9, France

Jean-Yves Gac

Office de la Recherche Scientifique et Technique Outre-Mer,
Centre de Dakar, BP 1386, Dakar, Sénégal

The great variability of rainfall in the Sahelian areas of Africa is well documented for all time scales during the Quaternary and historical periods. Since the beginning of the twentieth century the variability can be quantified from many meteorological and hydrological surveys[1–6]. The River Sénégal runoff gives a simplified view of the yearly mean rainfall over an extensive Sahelian and Sudanian area. The pattern of the annual river discharge over the nineteenth century shows short periods of drastic minimum followed by a slow return to longer periods of wet conditions with a double maximum. The mean time between the last three droughts is 31.3 yr with extreme conditions every 10.3 ± 4 yr. Extrapolating the curve of the mean annual modules suggests that the present drought should end in 1985 with full wet conditions being re-established in about 1992. If the same pattern continues, it is feared that a severe drought will occur around 2005.

In southern Sahara and northern Sahel the study of sediments, diatoms and pollens on cross-sections or cores of lacustrine deposits, has demonstrated that important environmental changes are frequent and repeated[7–12]. This variability of the climate during the past 40,000 yr is one of the main features of the Quaternary in this part of Africa[13]. On another time scale, Nicholson[14] and Maley[15] have shown that there have been alternating wet and dry periods during the past 1,000 yr. The Lake Chad level reconstitution (Fig. 1) shows how the apparent period of fluctuations depends on the method of data collection. The resolution is related to the accuracy and length of time

0028-0836/81/240475—04$01.00

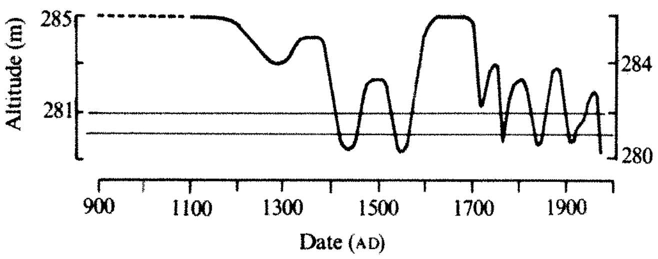


Fig. 1 Lake Chad level variations for the past 1,000 yr (from ref. 15). This curve synthesizes historical, geological and palynological data. Only the most important droughts have enough geological, historical or sociological impact to be recorded or memorized.

considered. On a 10^3- and 10^4-yr time scale, only important droughts have enough geological, historical or sociological impact to merit being recorded. This is why drought recurrence seems to be more frequent in recent times. Major famines and droughts during the past 300 yr, have been correlated by Nicholson[14] who found they repeated at intervals of 72 ± 16 yr (56–86 yr between the maxima). For both the nineteenth and twentieth centuries it seems that one drought could occur during every generation. For example, a rainfall maximum was experienced around 1841 when the River Sénégal flooded into the Ferlo valley as far up as Linguère[16,17]. A minimum rainfall appeared soon after, in 1848, when the Trarza from Mauritania crossed the River Sénégal to invade the Cayor and Oualo region[18], at about this time the Lake Chad level was low[15]. The years 1870–1900 were relatively humid[14]. They could

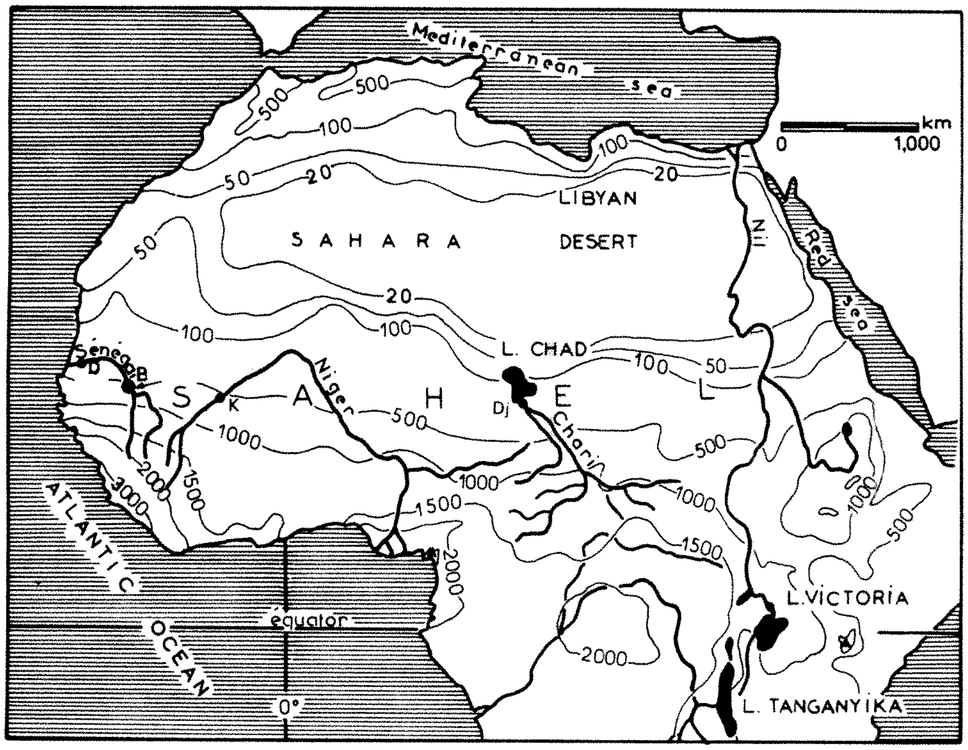


Fig. 2 Map showing the four main rivers flowing in Sahel. Isohyetes are from refs 28, 29. B, Bakel; D, Dagana; K, Koulikoro; Dj, Djamena.

Table 1 Mean annual runoff and maximum height of the River Sénégal at Bakel

Year	Runoff (m^3s^{-1})	H_{max} (m)	Year	Runoff (m^3s^{-1})	H_{max} (m)	Year	Runoff (m^3s^{-1})	H_{max} (m)	Year	Runoff (m^3s^{-1})	H_{max} (m)
1903	631	9.49	1923	755	11.05	1943	666	9.75	1963	666	10.11
1904	735	11.15	1924	1,244	12.20	1944	330	6.95	1964	970	12.56
1905	871	10.19	1925	839	11.00	1945	945	12.26	1965	1,048	12.48
1906	1,229	13.28	1926	520	8.00	1946	745	10.85	1966	841	11.70
1907	521	8.90	1927	1,074	12.25	1947	666	10.75	1967	1,037	11.94
1908	765	10.60	1928	903	11.70	1948	572	9.90	1968	397	8.94
1909	900	11.70	1929	898	11.70	1949	467	10.10	1969	764	10.11
1910	669	10.20	1930	837	11.00	1950	1,152	12.71	1970	542	9.70
1911	537	9.55	1931	738	10.70	1951	842	11.60	1971	598	10.72
1912	563	9.50	1932	768	11.20	1952	718	11.38	1972	263	6.24
1913	271	5.20	1933	831	11.70	1953	631	10.51	1973	361	8.38
1914	443	7.25	1934	699	11.60	1954	1,068	12.32	1974	760	11.91
1915	591	9.30	1934	1,164	12.35	1955	1,049	11.54	1975	602	10.14
1916	689	10.60	1936	1,234	12.70	1956	952	12.06	1976	470	6.96
1917	646	11.30	1937	644	9.90	1957	1,029	11.82	1977	324	7.06
1918	1,142	12.60	1938	807	11.80	1958	1,037	12.89	1978	523	7.87
1919	529	9.85	1939	559	9.65	1959	788	11.68	1979	301	6.20
1920	833	11.80	1940	430	8.75	1960	621	9.84	1980		8.66
1921	430	8.90	1941	417	8.95	1961	944	12.51	1981		
1922	1,218	13.19	1942	437	9.56	1962	769	10.80	1982		

Refs are given in Fig. 4 legend. The floodheight (H_{max}) is measured, after 1950, with an error of ±1 cm. The runoff is then better than 5% accuracy. Before 1950 the error could possibly have a maximum value of 30 cm, the accuracy then fell from 5% for high values of runoff to 15% for low values (see ref. 4). From the measured maximum flood in Bakel, 8.66 m on 11 September 1980, it can be estimated that the hydrological year 1980 (ending in April 1981) will have a runoff of more than 500 m^3s^{-1}, confirming the increase of 1976 and 1977.

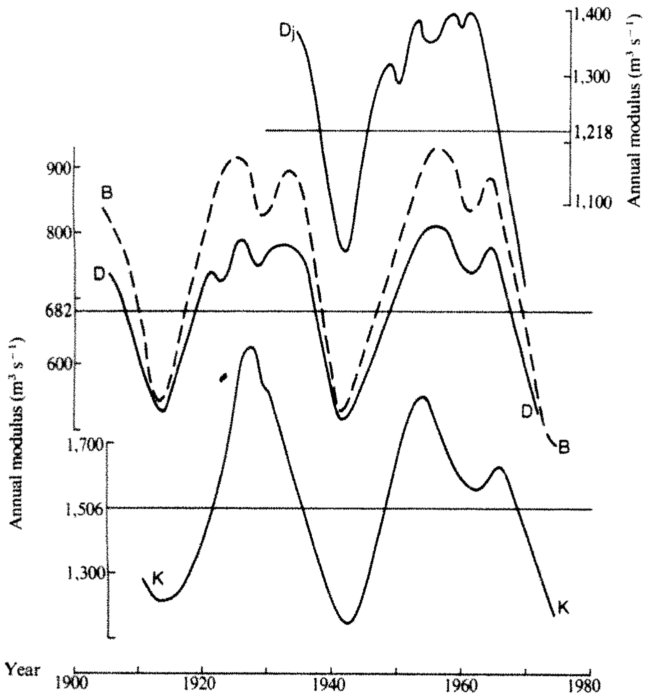


Fig. 3 Comparison of the pattern of river runoff in Sahel (7-yr running mean): Sénégal at Bakel (B) and Dagana (D), Niger at Koulikoro (K) and Chari at Djamena (Dj).

cover two periods of wet conditions because a break corresponding to a relative minimum is seen in Sénégal a short time before 1880 (ref. 14, Table 8.8).

Since the beginning of the twentieth century regular meteorological and hydrological surveys[2–6,19] have given a more quantitative and continuous set of data which illustrate the variability pattern of rainfall. The major rivers that flow through Sahel (Fig. 2), such as the Sénégal, the Niger, the Chari and partly the Nile, reflect the general annual and seasonal rainfall over a vast area of the Sahelian and Sudanian climatic zones. The annual discharges have been measured at several stations along the water course since the beginning of this century. The variation of the annual discharge shows a similar pattern for the various rivers[2] (Fig. 3): three drastic droughts clearly separated by two longer periods of humid conditions showing fluctuations mainly induced by flood plains evaporation, shape, lithology, soil, and so on. The River Sénégal has been taken as a typical example, and the Bakel hydrological station chosen because it is upstream from the main flood plains which experience great evaporation. The Bakel station (Table 1) gives a general idea of the mean rainfall over Sahelian and Sudanian zones. In both zones rainfall is a consequence of the North–South shifting of the monsoon. Meteorological data show that the annual rainfall in Sahel and Sudan fluctuates in the same direction[20,21].

To attenuate the effect of short-term variations, apparently 2 and 2.5 yr periods, curves have been constructed with various running means. We use the 7-yr running mean (Fig. 4) because the high and low points of this curve show the best coincidence with the years of maximum and minimum aridity. But similar results would be obtained with a 5-yr running mean more frequently used by hydrologists[5,19]. The curve (Fig. 4) shows several repeated characteristics. The slopes of the subvertical branches indicate that the onset of a drought is more rapid than the return to a wetter condition. Furthermore, each wet stage has a double maximum. This means that favourable conditions last ~18 yr, that is twice as long as arid periods characterized by peaks of ~8–12 yr. The drought maximum seems to recur after 31 ± 3 yr. If we consider only major events differing from the mean by more than 20%, the past 76 yr have seen seven of these extreme (maxima or minima) situations. They are separated by 8–15 yr; mean recurrence time is 10.3 yr. During the same period the solar cycle is 10.4 yr. Although the effect of solar cycles on climate is very controversial[22], indirect and complex relations could result from high atmosphere ionization[23,24] or geomagnetic field influence from solar wind[25,26].

The period of measured data (76 yr) is too short to establish a firm cyclicity in the drought appearance. Neither the nineteenth century observations, nor the records of the past 300 yr deny that a more or less dramatic drought could return at every generation, as claimed by the salt and trona producers whose livelihood depends on water table fluctuations[27]. If such a pattern continues some characteristics of the curve (Fig. 4) can be used to predict a date for the next rainfall maximum. The slope of the discharge curve during the transition to wet conditions can be used to estimate a slope between 14 and 27° for the next transition (Fig. 4). Consequently the long term average runoff should be reached at ~1985 (± 2 yr) and the return to a maximum wetness at around 1992 (±4 yr). However, this favourable period will undoubtedly end and a new drought will occur towards the beginning of the next millenary; at ~2005.

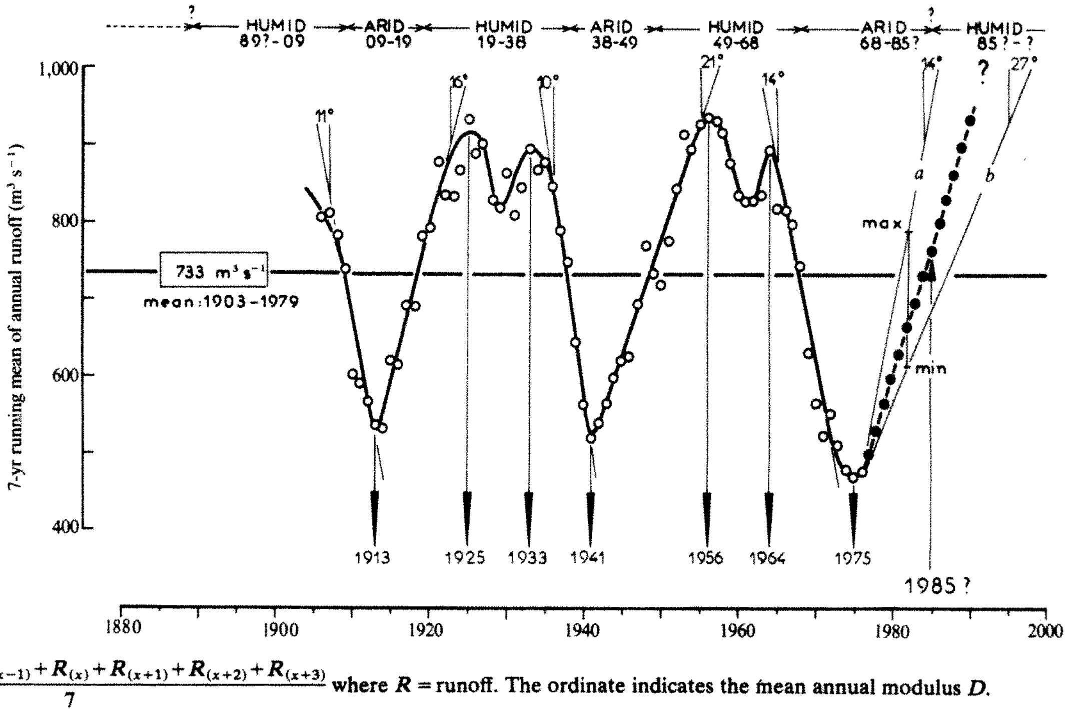


Fig. 4 Mean annual modulus of River Sénégal runoff measured at Bakel (7-yr running mean). The annual discharges have been compiled from documents of Senegal-Consult (1903–68), ORSTOM (1903–64), M. Juton (1969–71), A. Hamdinou (1971–75), ORSTOM (1975–78), Senegal Hydrological Service (1978–80). For predictive part of the curve (1977–92) *a* has the maximum slope (14°) of the preceding humid/arid change (the return to humid is slower than the humid/arid transition); the slope of *b* is given by deviation from the slopes of the two preceding fluctuations. The projected points are the mean annual modulus

$$D_{(\text{year}\,x)}=\frac{R_{(x-3)}+R_{(x-2)}+R_{(x-1)}+R_{(x)}+R_{(x+1)}+R_{(x+2)}+R_{(x+3)}}{7}$$

where R = runoff. The ordinate indicates the mean annual modulus D.

We thank D. R. Grant, L. Faure, J. J. Blanc, J. Fabre, F. Gasse, Ch. Toupet, M. Germain, Pr. Willis, O. Fambitakoye and N. A. Mörner for interesting comments or discussions. J. M. Grondin with Association Scientifique pour l'Etude du Quaternaire Africain (Dakar) colleagues in Dakar helped in the field and in bibliographical research. In addition to programmes of Office de la Recherche Scientifique et Technique Outre-Mer and Centre National de la Recherche Scientifique, Laboratoire de Géologie du Quaternaire, this work was stimulated by Délégation Générale à la Recherche Scientifique et Technique: action "Lutte contre l'aridité en milieu tropical" (Aide n° 78 7 0072), by International Geological Correlation Program projects n° 61 (Sea Level) and n° 146 (Lake Level and River Flood), and by International Union for Quaternary Research Feasibility Committee recommendations.

Received 11 December 1980; accepted 16 March 1981.

1. *Bibliographie Hydrogéologique de l'Afrique* (1968–1978) (Docum. Centre Etudes Géol. Minières, Paris, 1980).
2. *Monographie Hydrologique du Fleuve Niger* (ORSTOM, Paris, 1970).
3. *Monographie du Lac Tchad* (ORSTOM, Paris, 1969).
4. *Débit de Certains Cours d'Eau du Monde* Vol. II (UNESCO, Paris, 1971).
5. Rochette, C. *Monographie Hydrologique*, 1 (ORSTOM, Paris, 1974).
6. Rodier, J. A. *Travaux et Documents*, 46 (ORSTOM, Paris, 1975).
7. Servant, M. thesis (ORSTOM, Paris, 1973).
8. Servant-Vildary, S. *Travaux et Documents*, 84 (ORSTOM, Paris, 1978).
9. Gasse, F. *Nature* **265**, 42–45 (1977).
10. Livingstone, D. A. *A. Rev. Ecol. Syst.* **6**, 249–280 (1975).
11. Gac, J. Y. *Travaux et Documents*, 123, 1–251 (ORSTOM, Paris, 1980).
12. Street, F. A. & Grove, A. T. *Nature* **261**, 385–390 (1976).
13. Williams, M. A. J. & Faure, H. in *The Sahara and the Nile* (eds Williams, M. A. J. & Faure, H.) 583–586, (Balkema, Rotterdam, 1980).
14. Nicholson, S. E. in *The Sahara and the Nile* (eds Williams, M. A. J. & Faure, H.) 173–200 (Balkema, Rotterdam, 1980).
15. Maley, J. *Travaux et Documents* (ORSTOM, Paris, in the press).
16. Audiger, J. *Bull. Miss. Aménag. Sénégal* **112**, 1–139 (1957).
17. Monteillet, J., Faure, H., Pirazzoli, P. A. & Ravise, A. in *Paleoecology of Africa*, 13 (Balkema, Rotterdam, in the press).
18. Deschamps, H. *Le Sénégal et la Gambie* (Presses Universitaires de France, Paris, 1968).
19. Sircoulon, J. *Cah. O.R.S.T.O.M., Hydrol.* **13**, 75–174 (1976).
20. Toupet, Ch. & Michel, P. *Geol. Ecol Trop.* **3**, 137–157 (1979).
21. Grove, A. T. *Geogr. J.* **144**, 409 (1978).
22. Pittock, A. B. *Rev. Geophys. Space Phys.* **16**, 400–420 (1978).
23. Stuiver, M. & Quay, P. D. *Science* **207**, 11–19 (1980).
24. Eddy, J. A. *Climatic Change* Vol. 1, 173–190 (Reidel, Dordrecht, 1977).
25. Siscoe, G. L. *Nature* **276**, 348–352 (1978).
26. Fairbridge R. W. & Hillaire-Marcel, Cl. *Nature* **268**, 413–416, (1977).
27. Faure, H. *Rap. Dak.* 65–A7 *B.R.G.M. Rép. Niger*, 160 (1965).
28. Baumgartner, A. & Reichel, E. in *The World Water Balance*, 1–179 (Elsevier, Amsterdam, 1975).
29. Dubief, J. *Le climat du Sahara*, Vol. 2 (Institut de Recherches Sahariennes, Alger, 1963).

Geological evidence against major displacement in the Nares Strait

R. L. Christie*, P. R. Dawes†, T. Frisch‡, A. K. Higgins†, J. M. Hurst†, J. Wm. Kerr*§ & J. S. Peel†

* Geological Survey of Canada, Calgary, Canada T2L 2A7
† Grønlands Geologiske Undersøgelse, Øster Voldgade 10, 1350 København K, Denmark
‡ Geological Survey of Canada, Ottawa, Canada K1A OE8

Early continental drift[1–3] and later plate tectonic reconstructions[4–14] indicate left lateral (sinistral) movement along Nares Strait of between 150 and 400 km during the Cretaceous–Tertiary. A basic concept in such reconstructions is the substantial part played by seafloor spreading in the origin of Labrador Sea–Baffin Bay. Until recently, geological knowledge of the area adjacent to Nares Strait was insufficient to put tight constraints on displacement along Nares Strait (Fig. 1). Of seven independent lines of evidence from bedrock geology, termed markers, all are consistent with sinistral net displacement of Greenland along Nares Strait of 0–25 km. The three most imprecise markers can be reconciled with maximum movement of 100 km. This refutes conventional plate tectonic reconstructions which suggest movement of 150 km or more. New mechanisms should now be considered to explain the broad area of oceanic crust in Baffin Bay.

The foundation of the continental crust is the Precambrian shield that crops out in much of the southern part of the Nares Strait region (Fig. 1)[15–19]. In both Ellesmere Island and Greenland the shield consists of a high grade gneissic and metasedimentary terrain; a distinctive suite of acidic to basic igneous rocks has locally been transformed by Hudsonian deformation into complexly folded orthogneisses. Both regions contain marble-rich metasedimentary rocks; in Greenland these are mapped in the north as continuous linear, steeply inclined tracts up to 2 km thick that strike south-west near the coast of Nares Strait (Fig. 2). Similar marbles occur throughout the nunatak terrain of southeastern Ellesmere Island and these also form a metasedimentary-rich zone.

The marble-rich tracts in Ellesmere Island and Greenland (marker 1) are inferred to be part of the same supracrustal belt. No similar belt is known in the crystalline terrain elsewhere in the region of eastern Devon Island and northern Baffin Island in Canada or in the Thule and Melville Bugt areas in Greenland. This lithological continuity is evidence for correlating the shield areas of Ellesmere Island and adjacent Greenland directly across Nares Strait. No displacement thus need be invoked, and

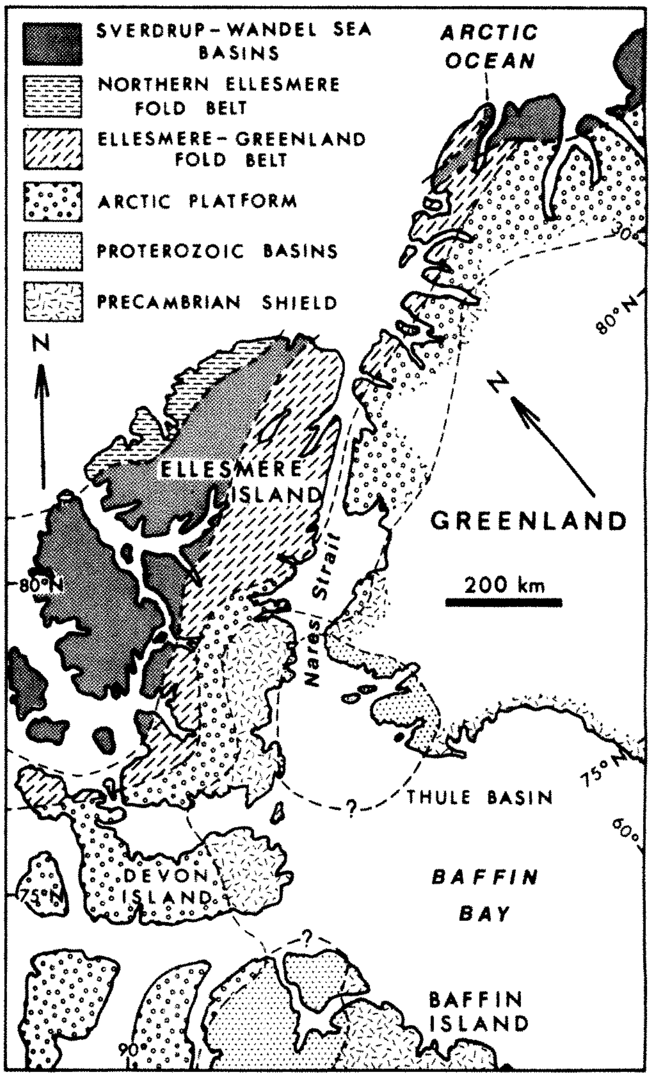


Fig. 1 The setting and main geographical features of the Nares Strait region together with the main geological provinces.

---

§ Present address: J. William Kerr & Associates Ltd, Suite 407, 630–8th Avenue S. W., Calgary, Alberta, Canada T2P 1G6.

0028–0836/81/240478—03$01.00

although this is an imprecise marker, sinistral displacement in excess of 100 km is considered unlikely.

The Thule Basin (marker 2) is an intracratonic basin spanning southernmost Nares Strait (Fig. 2) and consists of several formations of Proterozoic sedimentary and volcanic rocks, making up the Thule Group in north-west Greenland[18–21]. Some of the same formations occur on Ellesmere Island[15–17]. The Thule Group is unmetamorphosed and overlies the Precambrian Shield with profound unconformity. In Greenland the Thule Group is several kilometres thick; it consists essentially of a lower sequence of clastic sediments with basalt flows and sills, and an upper sequence of sandstones, shales and carbonate rocks. Only the lower sequence is present on Ellesmere Island[17]. Unit to unit matching, with excellent marker beds, allows a reliable correlation between Ellesmere Island and Greenland[17,21]. The strata thin northwards onto the Bache Peninsula Arch both in Greenland and Ellesmere Island, defining the northern margin of the Thule Basin. No apparent displacement of this margin occurs across Nares Strait[22–24], but sinistral movements up to ~50 km could be accommodated.

The Bache Peninsula Arch[25] (marker 3) is a structural feature composed of Precambrian Shield rocks that must have been relatively high structurally, at least as early as late Proterozoic time. Thule Group strata over the arch in both Ellesmere Island and Greenland have a remarkable similarity in thickness and lithology, as illustrated by the classic platform sequences on Bache Peninsula and in Inglefield Land[26–28]. The Bache Peninsula Arch is also reflected in thicknesses of Cambrian to Ordovician rocks in Ellesmere Island and Greenland. The arch is a broad feature, not narrowly delineated on either side of Nares Strait, but its present trend across the strait is consistent with its having had the same trend in at least Proterozoic and early Palaeozoic time. However, as it is an imprecise marker some displacement could be accommodated, though it cannot have been more than 100 km in a sinistral sense.

The Franklinian Geosyncline is a linear basin that extends northeastwards through the Canadian Arctic Islands and northern Greenland[18,29–31]. The geosyncline contains a thick column of Upper Proterozoic to Upper Devonian sediments. Several linear features of the Franklinian Geosyncline (markers 4–6) are pertinent to displacement along Nares Strait.

The Franklinian Geosyncline passes southeastwards into equivalent rocks of the Arctic Platform. The boundary separating them is a hinge-line, across which sedimentary thicknesses increase rapidly northwards. The hinge-line migrated laterally from time to time on Ellesmere Island[31] and for the most part it coincides with the south-east boundary of late Palaeozoic folding (marker 4, Fig. 2). This marker is a structural boundary which might have an enhanced value because it generally coincides with the hinge-line that existed before the strait formed.

The south-east limit of folding (marker 4) is well delineated both on Ellesmere Island[29,31] and in Greenland[18,24]. The limit of folding has an apparent offset along Nares Strait and this has been used[32,33] to suggest that there was at least 250 km of left lateral strike-slip displacement. It was suggested[32,33] that the southernmost folds of Ellesmere Island originally continued directly across the strait to the southernmost folds in Greenland, and that the present apparent offset was created by left lateral strike-slip faulting after the folding. However, this is not substantiated by the detailed structural geology which shows that in eastern Ellesmere Island, close to Nares Strait, fold trends swing markedly from easterly to northeasterly[29–31,34–36]. Trends of individual folds run parallel to Nares Strait for some distance. Similar swings in fold trends occur in northern Greenland[18,24], and farther west in Ellesmere Island[29–31]. The swings may coincide with salients and re-entrants in the hinge-line at the margin of the original sedimentary basin. We believe that the south-east limit of folding reflected the change in fold trends in easternmost Ellesmere Island, and had an original bend from Ellesmere Island northwards along the line of Nares Strait itself (Fig. 2), and then bent again eastwards into Greenland. Marker 4 exhibits its closest match across Nares Strait with models of displacement between 0 and 50 km, while displacements above 50 km create unacceptable divergences in fold trends.

Two important Lower Palaeozoic facies boundaries occur on Ellesmere Island and in Greenland. Marker 5 represents the Cambrian to Upper Ordovician trough margin, and marker 6 the Lower Silurian to Devonian facies front between platform carbonates and mudstones. The Cambro-Ordovician facies boundary on Ellesmere Island (Fig. 2) separates thick, shallow water carbonates of the miogeocline to the south-west from equivalent mudstones, resedimented conglomerates, cherts and turbidites of the Hazen Trough farther north-west[37]. This marker is linear and a similar feature in eastern North Greenland is the site of faults[38,39]. Facies relations on northern Hall Land and Nyeboe Land together with the linear nature of the coastline suggest that this is the site of marker 5 in western North Greenland. The linear boundary in Ellesmere Island is on direct strike with the similar boundary in northern Greenland, and suggests that only small displacements can have taken place; displacement in excess of 50 km can probably be excluded.

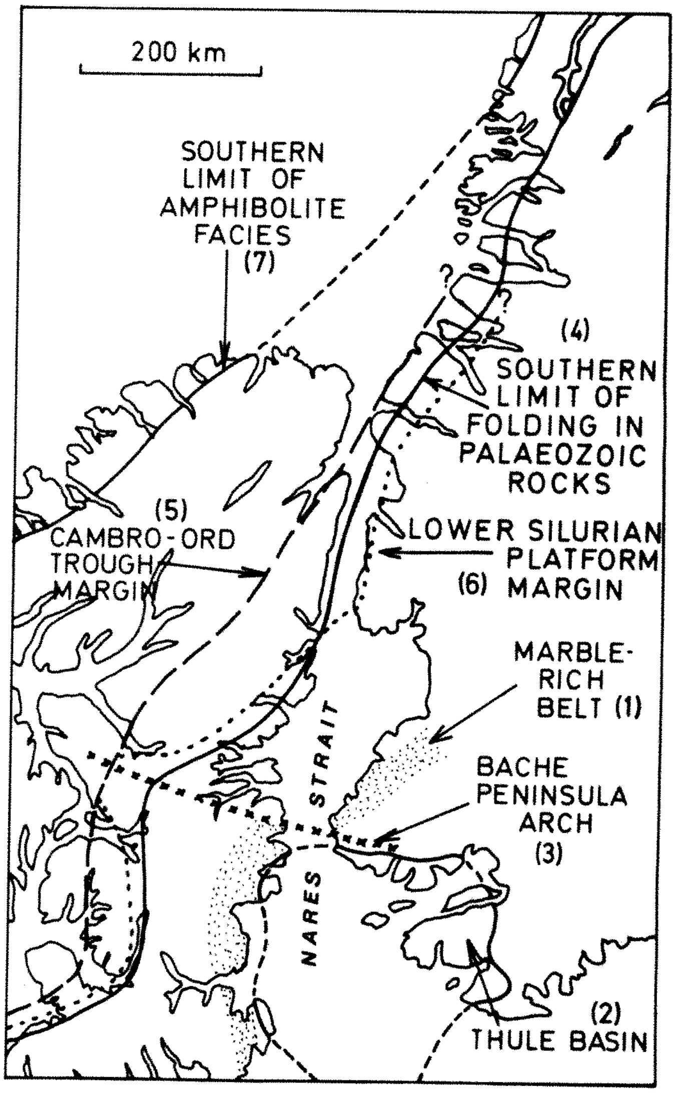

200 km
SOUTHERN LIMIT OF AMPHIBOLITE FACIES (7)
?
(4) SOUTHERN LIMIT OF FOLDING IN PALAEOZOIC ROCKS
(5) CAMBRO-ORD TROUGH MARGIN
LOWER SILURIAN PLATFORM (6) MARGIN
MARBLE-RICH BELT (1)
NARES STRAIT
BACHE PENINSULA ARCH (3)
(2) THULE BASIN

Fig. 2 Geological 'markers' that occur on both sides of Nares Strait that are useful for assessing displacement.

The Lower Silurian platform margin (marker 6) represents the facies front between the clastics of the Hazen Trough and the carbonates of the Franklinian Miogeocline and has a position to the south-east of marker 5. The location of this marker is shown in Fig. 2 in Greenland[40–42] and Ellesmere Island[43] during Middle Llandovery time. A feature of the position of this boundary with time is its southerly migration. The diachronous southerly spread of the mudstone facies fringing the Hazen Trough started during the latest Ordovician in Ellesmere Island and reached western North Greenland in the Early Silurian[41,42] indicating facies continuity across Nares Strait. The facies boundary as depicted in Fig. 2 lasted until Late Silurian time in Greenland and until Devonian time on Ellesmere Island. This is perhaps the best of all the markers in that it is a precisely dated, discrete, vertical boundary and crosses Nares Strait at one of its narrowest points. It strikes across the strait without offset, but because of the meandering nature of facies fronts sinistral displacement of up to 25 km cannot be ruled out.

Fold and metamorphic intensity increases northwards in Ellesmere Island and northern Greenland. The Pearya Geanticline, a linear tectonic feature cored by metamorphic and igneous rocks, trends northeastwards on northern Ellesmere Island. The events in the long tectonic history of the Pearya Geanticline are well established[30]. A region of high-grade amphibolite facies metamorphism (Fig. 2) occurs in the geanticline on northern Ellesmere Island[44,45], and the southeastern margin of amphibolite facies rocks is given here as marker 7. Amphibolite facies rocks also occur in northern Greenland in a narrow, intensely deformed zone in the extreme north[46] where their southeastern margin is regarded as marker 7. The amphibolite-grade metamorphic belts of northern Ellesmere Island and North Greenland are fundamental, deep rooted features, a type which is not prone to rapid changes in trend. If the original boundary was approximately straight, which seems reasonable, the present situation of the amphibolite facies rocks (Fig. 2) is consistent with minor displacement along Nares Strait. However, the two regions are separated by a wide expanse of water and displacements of up to 100 km could be accommodated with approximately linear connections between the two areas; models with greater displacement become increasingly tenuous.

Each of the seven markers has some bearing on the amount of maximum net displacement that can have occurred along Nares Strait, the amount varying with the nature and the reliability of the geological data defining the marker. Three imprecise markers could accommodate up to 100 km displacement, while three other markers could be reconciled with up to a maximum of 50 km sinistral movement. All the markers are consistent with net sinistral displacement in the range 0–25 km; marker 6, which permits the narrowest range of possible movement, is also considered to be the most precise of the markers. We therefore conclude that net strike-slip displacement along Nares Strait was minor, and not greater than 25 km.

Some of the above markers have been used[32,33] to support displacements of up to 300 km along Nares Strait. The correlation of the south-west limit of folding (marker 4) directly across the strait has been mentioned above. The Thule Basin deposits of Greenland (marker 3) have been correlated in such models with Proterozoic deposits of northern Baffin Island (Fig. 1). There is a general similarity between these strata, and also strata in other Proterozoic basins on the northern margin of the North American craton; however, models with such large displacements violate the detailed correlation of the Thule Basin rocks in Greenland and Ellesmere Island. Correlation of faults and thrusts on the two sides of Nares Strait have also been suggested in models with large displacements, but some of these structures are associated with important Tertiary movements[25,31] and are thus of little value in determining the position of land masses in pre-drift reconstructions. Faults and thrusts in Greenland can be correlated with other, on-line, dislocations in Ellesmere Island without invoking displacement.

The linearity and length of Nares Strait does, however, suggest that it is the site of a movement zone. Indeed, we believe that motion probably has occurred along it, particularly in view of the fault zone of Judge Daly Promontory[25,47–49] and the suggested origin of the ocean basins surrounding the region. Left lateral strike-slip movement of 19 km on the fault zone is estimated by Mayr[49] from offset of limbs of an intersected fold.

The Nares Strait is an important element in plate tectonic reconstructions of the Northern Hemisphere. Geological data placing constraints on the amount of net movement that may have occurred along the strait have wide geological implications[50]. Thus, without invoking a complicated movement history for Nares Strait (one might speculate dextral displacement(s) in the late Palaeozoic or/and Mesozoic time followed by sinistral displacement(s) associated with the formation of the surrounding oceans) the geological evidence conflicts with the widely accepted origin of Baffin Bay[8,12,32,51–53], in which lateral movement by seafloor spreading approximates the distance separating the present continental shelves. Rigid-plate models have failed to produce an all embracing solution for the origin of Baffin Bay (and also Labrador Sea) that takes into account the geological restraints of the lands bordering the oceans. We would like to see a closer coordination of the onshore geological data into the discussion of the origin of the surrounding oceans that are mainly based on offshore geophysical data.

P.R.D., A.K.H., J.M.H. and J.S.P. publish with the permission of the Director of the Geological Survey of Greenland. R.L.C., T.F. and J.W.K. thank the Director of the Geological Survey of Canada for permission to publish this paper.

Received 19 January; accepted 14 April 1981.

1. Taylor, F. B. *Bull. geol. Soc. Am.* **21,** 179–226 (1910).
2. Wegener, A. *Die Entstehung der Kontinente und Ozeane* (Vieweg, Braunschweig, 1922).
3. Carey, S. W. in *Continental Drift—A Symposium* (ed. Carey, S. W.) 177–355 (Geology Department, University of Tasmania, Hobart, 1958).
4. Wilson, J. T. *Scient. Am.* **208,** 86–100 (1963).
5. Wilson, J. T. *Phil. Trans. R. Soc.* **A258,** 145–167 (1965).
6. Bullard, E. *et al. Phil. Trans. R. Soc.* **A258,** 41–51 (1965).
7. Dewey, J. F. *Scient. Am.* **226,** 56–68 (1972).
8. Keen, C. E. *et al. Can. J. Earth Sci.* **9,** 239–256 (1972).
9. Pitman, W. C. & Talwani, M. *Bull. Geol. Soc. Am.* **83,** 619–646 (1972).
10. Harland, W. B. *Mem. Am. Ass. petrol. Geol.* **19,** 599–608 (1973).
11. Beh, R. L. *Mem. Can. Soc. petrol. Geol.* **4,** 453–476 (1975).
12. Strivastava, S. P. *Geophys. J. R. astr. Soc.* **52,** 313–357 (1978).
13. Irving, E. *Can. J. Earth Sci.* **16,** 669–694 (1979).
14. Kristoffersen, Y. & Talwani, M. *Bull. Geol. Soc. Am.* **88,** 1037–1049 (1977).
15. Christie, R. L. *Geol. Surv. Can. Map* **9** (1962).
16. Christie, R. L. *Geol. Surv. Can. Map* **12** (1962).
17. Frisch, T. *et al. Geol. Surv. Can. Pap.* **78–1A,** 135–138 (1978).
18. Dawes, P. R. in *Geology of Greenland* (eds Escher, A. & Watt, W. S.) 248–303 (Geological Survey of Greenland, Copenhagen, 1976).
19. Davies, W. E. *et al. Meddr Grønland* **162,** (12) 1–62 (1963).
20. Dawes, P. R. *Rapp. Grønl. geol. Unders.* **80,** 23–28 (1976).
21. Dawes, P. R. *Rapp. Grønl. geol. Unders.* **95,** 14–22 (1979).
22. Christie, R. L. *et al. Abstr. Prog. Geol. Ass. Can., Min. Ass. Can. Geol. Soc. Am.* **3,** 380 (1978).
23. Dawes, P. R. *Norsk Polarinst. Skrifter* **167,** 321–324 (1979).
24. Dawes, P. R. & Peel, J. S. in *Ocean Basins and Margins,* Vol. 5 (eds Nairn, A. E. M. *et al.*) (Plenum, New York, 1980).
25. Kerr, J. Wm. *Bull. Can. Petrol. Geol.* **15,** 483–520 (1967).
26. Troelsen, J. C. *Meddr Grønland* **149,** (2) 1–85 (1950).
27. Cowie, J. W. *Meddr Grønland* **164,** (3) 1–47 (1961).
28. Christie, R. L. *Mem. geol. Surv. Can.* **347,** 1-61 (1967).
29. Thorsteinsson, R. & Tozer, E. T. *Geol. Surv. Can. econ. geol. Rep.* **1,** 548–590 (1970).
30. Trettin, H. P. *et al. Geol. Ass. Can. Spec. Pap.* **11,** 83–179 (1972).
31. Kerr, J. Wm. in *Ocean Basins and Margins* Vol. 5 (eds Nairn, A. E. M. *et al.*) (Plenum, New York, 1980).
32. Newman, P. H. thesis, Dalhousie Univ. (1977).
33. Newman, P. H. & Falconer, R. K. H. *Abstr. Prog. geol. Ass. Can., Min. Ass. Can. geol. Soc. Am.* **3,** 463 (1978).
34. Kerr, J. Wm. *Geol. Surv. Can. Map* **1357A** (1973).
35. Kerr, J. Wm. *Geol. Surv. Can. Map* **1358A–1** (1973).
36. Kerr, J. Wm. *Geol. Surv. Can. Map* **1358A–2** (1973).
37. Trettin, H. P. *Bull. geol. Surv. Can.* **272,** 1–84 (1979).
38. Surlyk, F. *et al. Nature* **286,** 800–803 (1980).
39. Hurst, J. M. & Surlyk, F. *Rapp. Grønl. geol. Unders.* **99,** 73–78 (1980).
40. Hurst, J. M. *Rapp. Grønl. geol. Unders.* **88,** 41–49 (1980).
41. Hurst, J. M. *Bull. Am. Ass. petrol. geol.* **64,** 527–548 (1980).
42. Hurst, J. M. *Bull. Grønl. geol. Unders* **138,** 1–96 (1980).
43. Kerr, J. Wm. *Bull. geol. Surv. Can.* **260,** 1–55 (1976).
44. Frisch, T. *Bull. geol. Surv. Can.* **229,** 1–87 (1974).
45. Frisch, T. *Geol. Surv. Can. Pap.* **76–1A,** 429–430 (1976).
46. Dawes, P. R. & Soper, N. J. *Mem. Am. Ass. petrol. Geol.* **19,** 117–134 (1973).
47. Christie, R. L. *Mem. geol. Surv. Can.* **331,** 1–79 (1964).
48. Christie, R. L. *Geol. Surv. Can. Pap.* **74–1A,** 297–299 (1974).
49. Mayr, U. *Abstr. Prog. geol. Ass. Can., Min. Ass. Can.* **5,** 69 (1980).
50. Kerr, J. Wm. *Bull. Can. petrol. Geol.* **28,** 279–283 (1980).
51. Keen, C. E. & Hyndman, R. D. *Can. J. Earth Sci.* **16,** 712–747 (1979).
52. Ross, D. I. *et al. Geol. Surv. Can. Pap.* **71–23,** 633–638 (1973).
53. Jackson, H. R. *et al. Can. J. Earth Sci.* **16,** 2122–2135 (1979).

Molecular detail in electron micrographs of quaterrylene $C_{40}H_{20}$

David J. Smith

High Resolution Electron Microscope, University of Cambridge,
Free School Lane, Cambridge CB2 3RQ, UK

J. R. Fryer

Chemistry Department, University of Glasgow,
Glasgow G12 8QQ, UK

The electron microscopy of organic compounds is beset by the problem of radiation damage due to the electron beam[1]. This radiation sensitivity is somewhat reduced in aromatic systems[2] and organometallic complexes[3-5] but it, nevertheless, limits the fineness of detail which can be resolved in images from these types of materials. Previous studies of the more stable compounds have concentrated on those where molecular superimposition normally occurred on a major crystallographic axis. The compound quaterrylene, $C_{40}H_{20}$, does not have such a projection axis and is therefore more typical of most organic crystals. We report here results from the examination of quaterrylene[6] in the Cambridge University 600-kV high resolution electron microscope (HREM)[7,8]. These results provide direct information about intermolecular stacking and, after lattice averaging, show detail at the level of the six-membered aromatic rings within molecules of this material.

The limitations posed by radiation damage in the electron microscope can be partially circumvented by spatial averaging over many unit cells, the use of sensitive photographic emulsions such as X-ray film, and recording images from previously unexposed areas[1,2,9,10] (minimum exposure technique[11]). Clearly, spatial averaging does not provide direct information about nonperiodic features, although it permits structure determination from sample volumes many orders of magnitude less than other techniques such as X-ray diffraction and it was of particular value to the present study where the specimen consisted of very small domains (~10 nm)[3]. Furthermore, X-ray film has a superior modulation transfer function to that of normal electron microscope film[12,13] which is of particular relevance in studies of beam-sensitive specimens due to the very small electron exposures required to reach optimum optical density. Higher accelerating voltages have also been reported to reduce the rate of damage in some materials[14]. Although examination of this effect was not the purpose of the present work, the improved contrast transfer function (interpretable resolution) of the microscope used, relative to conventional 100-kV machines[15], should lead to an enhanced signal-to-noise ratio and hence improved object definition.

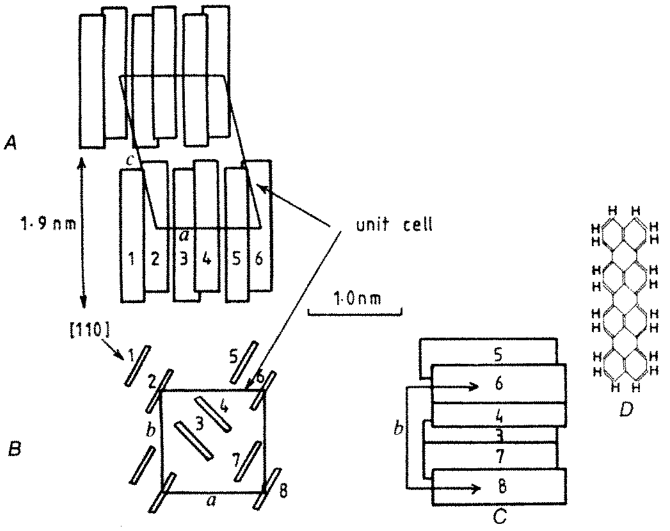


Fig. 1 Crystal and molecular structure of quaterrylene. Molecules are represented as rectangles in crystal projections, numbering referring to the same molecule in each projection. *A*, *b*-axis projection, two rows of molecules; *B*, end-on view of molecules in orthogonal *c*-axis projection; *C*, *a*-axis projection; *D*, molecular structure.

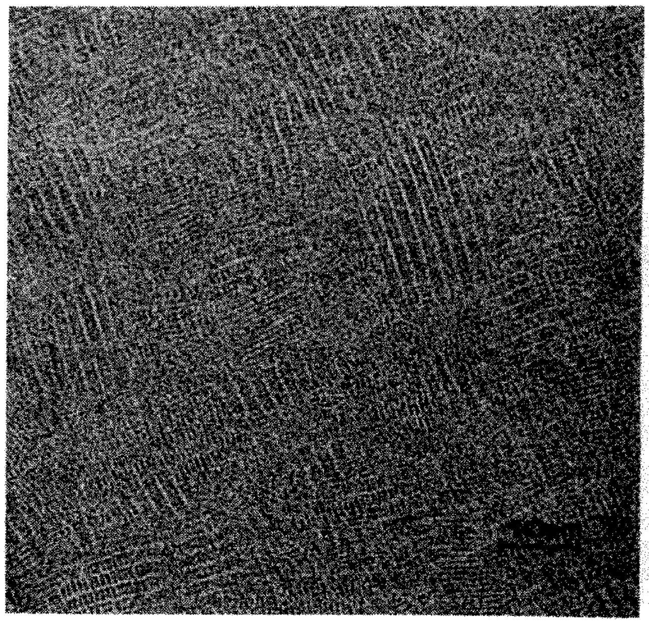


Fig. 2 Low-magnification micrograph of a crystalline mosaic of quaterrylene evaporated onto KCl (100).

Quaterrylene is a planar polynuclear aromatic hydrocarbon whose molecular structure is shown in Fig. 1*D*. The crystal structure[16] is monoclinic with $a = 1.115$, $b = 1.063$, $c = 1.924$ nm, $\beta = 100.46°$, $z = 4$ and space group $P2_{1/a}$. Projections of this structure are shown in Fig. 1*A*, *B*, *C* with the molecules represented by rectangles. These projections show that there was not exact molecular superimposition on any major axis, but that a {110} projection—despite its consisting of molecular pairs at different inclinations—may give rise to near superimposition of regions of high electron density (such as the naphthalene moieties) that would provide information about molecular positions and orientation.

Vacuum evaporation of quaterrylene onto a KCl (001) substrate provided a crystalline mosaic of <10 nm thickness (Fig. 2) with small crystals oriented around the (110) plane. This plane aligns approximately parallel to the substrate. Deviations from this orientation were within a few degrees and occurred randomly. The *c*-axis projection of the crystal structure (Fig. 1*b*) shows that images should be obtained of the molecular columns down a {110} axis, with molecules having their molecular planes slightly inclined, at 9° and 81°, to the direction of the incident electron beam.

Operation of the HREM for these studies has primarily been at an accelerating voltage of 500 kV, with magnifications typically of around ×100,000. Minimum exposure techniques were used, with a low-inertia beam shutter above the specimen[17] blanking off the beam while a fresh specimen area was shifted into place ready for photographic recording. The optimum defocus (~ −390 nm) was difficult to estimate experimentally because of the relatively low magnification, the low specimen contrast and the low beam intensity. Furthermore, there were often slight variations in vertical position across the specimen. Consequently, three exposures with focal differences of 75 nm were normally taken of each area of the specimen. Typical electron doses for an exposure were in the range 1,000–2,000 electrons nm^{-2}. This dose causes severe damage in aliphatic

0028-0836/81/240481—02$01.00

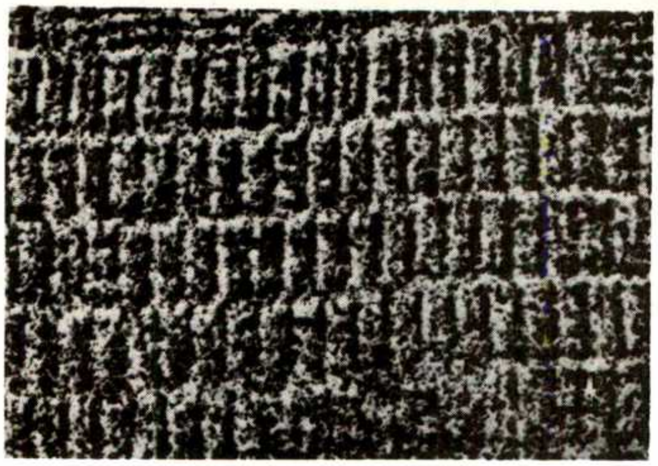


Fig. 3 Image of four rows of molecules in (110) projection. Slight variations in inclination down (110) giving rise to differences in contrast along each row and the curve in the 1.9-nm lattice spacing probably arising from the fusion of two crystals nucleated at different points on the KCl substrate.

systems[18] but for an aromatic hydrocarbon should not excessively limit the image resolution. Considerable blackening of the reverse side emulsion occurred due to penetration of the high-energy electron beam through the film backing, as well as back-scattered radiation from the film holder. As this low resolution information obscured the primary image, it was normally removed before serious analysis of image features was undertaken.

As a consequence of the very low electron dose used for recording, as well as the minimal exposure of the sample to the electron beam, our images revealed details of quaterrylene molecules at a level hitherto unattained in previous images of aromatic hydrocarbons in the electron microscope; optical diffractogram analysis of the micrographs indicated the transfer of specimen information to resolutions better than 0.35 nm. Figure 3 shows the projection of the molecular columns down a {110} axis whilst Fig. 4 shows structure that dimensionally corresponds to the molecular projection as drawn. (The signal-to-noise ratio of the latter image has been enhanced by a factor of ~3 by photographic superimposition 10 times.) The pattern of contrast observed can be related to the positions of the double and single rings of quaterrylene occurring within the (110) column, although the crystal structure of this material, as pointed out above, precludes the exact superimposition of molecules down this axis.

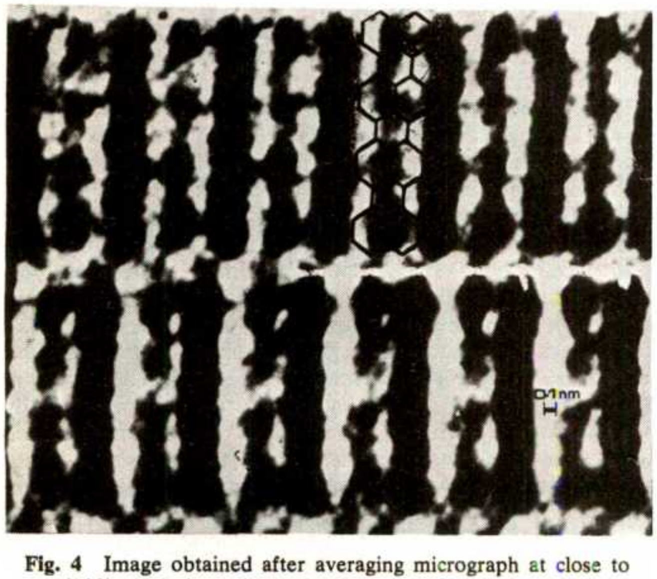


Fig. 4 Image obtained after averaging micrograph at close to (110) projection, with molecular structure superimposed.

We cannot yet judge whether the improved molecular detail in our micrographs is a direct consequence of the increased beam voltage, as no comparable study of quaterrylene has been undertaken at lower voltages but we hope soon to observe anthanthrene for which such information exists[2]. Certainly, quaterrylene is slightly more beam stable than many hydrocarbons, and is comparable with copper phthalocyanine and nucleic acid bases. For more sensitive materials some sort of low temperature specimen stage, as well as the other methods used here, may be necessary to resolve detail on the molecular scale[19,20].

The ability to resolve detail in images of aromatic hydrocarbons to better than 0.35 nm, would permit molecules of similar size to quaterrylene to be located and their relative configurations to be determined at grain boundaries and other nonperiodic regions of thin crystals—of particular relevance is the van der Waals distances of ~0.35 nm between molecules in these materials where no other intermolecular forces are present. An examination of thin film annealing, grain boundary structure and dislocations in quaterrylene has certainly proved possible—a full report will be published elsewhere. Such direct information about morphological features should prove essential to any complete explanation for a wide range of physical properties (such as the electrical conductivity in thin quaterrylene films) of these materials.

The Cambridge University 600-kV high resolution electron microscope has been built as a joint project between the Cavendish Laboratory and the Department of Engineering with major financial support from the SRC.

Received 2 February; accepted 1 May 1981.

1. Glaeser, R. M. in *Physical Aspects of Electron Microscopy and Microbeam Analysis* (eds Siegel, B. M. & Beaman, D. R.) 205–230 (Wiley, New York, 1975).
2. Fryer, J. R. *Acta crystallogr.* A**34**, 603–607 (1978).
3. Uyeda, N., Kobayashi, T., Suito, E., Harada, Y. & Watanabe, M. *J. appl. Phys.* **43**, 5181–5189 (1972).
4. Uyeda, N., Kobayashi, T., Ishizuka, K. & Fujiyoski, Y. *Nature* **285,** 95–97 (1980).
5. Murata, Y., Fryer, J. R. & Baird, T. *J. Microsc.* **108,** 261–275 (1976).
6. Fryer, J. R. & Smith, D. J. in *Proc. 37th EMSA Meet.* (ed. Bailey, G. W.) 552–553 (Claitor's, Baton Rouge, 1979).
7. Nixon, W. C. *et al.* in *Developments in Electron Microscopy and Analysis 1977* (ed. Missell, D. L.) 13–16 (Institute of Physics, Bristol, 1977).
8. Cosslett, V. E. *Proc. R. Soc.* A**370,** 1–16 (1980).
9. Kuo, I. A. M. & Glaeser, R. M. *Ultramicroscopy* **1,** 53–66 (1975).
10. Cosslett, V. E. *J. Microsc.* **113,** 129 (1978).
11. Williams, R. C. & Fisher, H. W. *J. molec. Biol.* **52,** 121–123 (1970).
12. Dorignac, D., Maclachlan, M. E. C. & Jouffrey, B. *Ultramicroscopy* **2,** 49–52 (1976).
13. Grano, D. A. & Downing, K. H. *Proc. 38th EMSA Meet.* (ed. Bailey, G. W.) 228 (Claitor's, Baton Rouge, 1980).
14. Glaeser, R. M. in *High Voltage Electron Microscopy* (eds Swann, P. R., Humphreys, C. J. & Goringe, M. J.) 370–378 (Academic, London, 1974).
15. Cosslett, V. E. *et al. Nature* **281,** 49–51 (1979).
16. Kerr, K. A., Ashmore, J. P. & Speakman, J. C. *Proc. R. Soc.* A**344,** 199–215 (1975).
17. Fryer, J. R., Cleaver, J. R. A. & Smith, D. J. in *Electron Microscopy and Analysis 1979* (ed. Mulvey, T.) 287–290 (Institute of Physics, Bristol, 1980).
18. Chiu, W. *Scanning Electron Microsc.* **1,** 569 (1978).
19. Dietrich, I., Formanek, H., Fox, F., Knapek, E. & Weyl, R. *Nature* **277,** 380 (1979).
20. Hayward, S. B. & Glaeser, R. M. *Ultramicroscopy* **4,** 201–210 (1979).

Model protocells photochemically reduce carbonate to organic carbon

Clair Folsome & Andrew Brittain

Department of Microbiology, University of Hawaii, Honolulu, Hawaii 96822, USA

Synthetic cell-sized organic microstructures effect the long-wavelength UV photosynthesis of organic products from carbonate. Formaldehyde is the most abundant photoproduct and water is the major proton donor for this reduced form of carbon. The apparent quantum yield is $\sim 10^{-5}$ carbon atoms per incident UV 254-nm photon. We show here that these results for model phase-bounded systems are consistent with the postulate that metabolism of progenitors to the earliest living cells could have been, at least in part, photosynthetic.

0028–0836/81/240482—03$01.00

A 'protocell' is here defined as an organic cell-sized phase-bounded structure which arose from organic precursors of abiological origin. We propose that phase-bounded structures of organic composition were probably progenitors to the first living cell, and that such protocells were in existence for perhaps 500–1,000 Myr (refs 1, 2).

Protocells have been the subject of numerous modelling studies[3-8] which have raised several questions, such as the plausibility of protocell formation in primordial conditions and the suitability of protocells to evolve towards biological forms. However, all protocell models mirror features of contemporary cells such as size and the presence of phase boundaries.

The existence of phase boundaries necessitates that there is a difference in chemical composition and energy level between protocell interior and the environment. Viewing metabolism as changes in concentrations and reaction rates of compounds across phase boundaries, then the beginnings of cellular metabolism coincide with the formation of protocells. Many alternative models for protocell metabolism can be adduced within these general constraints. We have designed a model system which demonstrates the photosynthesis of formaldehyde and other organic products by the long-wavelength UV irradiation of aqueous suspensions of organic model protocells and carbonate under a nitrogen/hydrogen atmosphere.

Phase-bounded organic microstructures were prepared from quenched spark discharges through gas mixtures of N_2 (50 mm), CH_4 (200 mm) and CO_2 (50 mm) on a distilled water surface as described elsewhere[1,8]. The structures' form can be disrupted by freeze-thawing. The osmotically sensitive nature of these microstructures is revealed by microscopic examination: salt equilibrated (1 M NaCl) suspensions which are shocked by sudden exposure to distilled water will swell and burst. Up to 60% of the total recovered organic carbon is phase-bounded. These microstructures display complex morphologies and occur in a range of sizes 1–20 μm in diameter. Their number increases autocatalytically during synthesis, which indicates that self-assembly of new structures on surfaces of pre-existing structures occurs[1].

Microstructures were washed by centrifugation in distilled water and stored frozen or used immediately. For all carbonate reduction experiments, septum-stoppered quartz tubes 12 × 100 mm were used containing 2 ml of microstructure suspension (about 7 mg dry wt microstructures) and a 6-ml headspace which was filled, after repeated flushing, with 710 mm N_2 and 50 mm H_2. A small Teflon-coated magnetic flea was used to stir the mixture continuously. ^{14}C-sodium bicarbonate at a specific activity of 50 Ci M^{-1} (Research Products International) was added and reaction mixtures were irradiated in a Rayonet photochemical reactor which contained 16 low-pressure mercury 8-W lamps with Vycor envelopes. UV 254-nm intensities were measured by thymine dimer actinometry using the same size quartz tubes[9] and were 249 erg s^{-1} mm^{-2}. The UV 254-nm radiation incident on exposed surfaces of the quartz reaction vessels was 49.9×10^3 erg s^{-1}. Temperature at the reaction vessel site was 27 °C. Unirradiated control experiments were wrapped in two layers of aluminum foil and placed in the photochemical reactor.

After 24 h exposure, the reaction vessel was evacuated, then acidified with 1.5 ml 11 M HCl, flushed with N_2 and degassed under vacuum for 30 min to remove unreacted carbonate. The sample was centrifuged at 12,000*g* for 5 min to yield a soluble and a particulate phase: the latter was rinsed several times and washings were combined with the soluble phase. Both particulate and soluble fractions were assayed for ^{14}C activity by liquid scintillation counting. The sum of these two fractions' activities is termed the total organic carbon produced in the reaction from labelled carbonate.

Generally 30% or less of the total recovered organic carbon was found in the particulate fraction: the balance was soluble organic carbon. These results are presented in Table 1. The data show that carbonate is photoreduced to organic carbon. The reaction requires the presence of organic microstructures and UV 254-nm radiation (see control experiments 1 and 2). Higher yields of reduced carbon are obtained if the microstructures are intact and not disrupted by freeze–thaw cycles (experiment 4). The apparent quantum yield (the number of carbon atoms reduced to organic carbon per incident UV 254 nm photon) is as high as 2.4×10^{-5}. The actual quantum yield of this reaction is higher, as no corrections have been made for photons actually adsorbed by the reaction mixtures.

The soluble organic fraction was concentrated by partial lyophilization and separated by thin-layer chromatography (TLC) on silica gel-H in various solvent systems. Bands 1 cm wide were scraped from the plates and ^{14}C activity measured by liquid scintillation counting. About 70% of the applied label was found to migrate with R_f values identical to those of an authentic ^{14}C-formaldehyde standard in the following solvent systems: (1) methanol/water-saturated $(NH_4)_2CO_3$, 3/1; (2) 2-butanol/water/formic acid, 17/2/1; (3) chloroform/methanol/4.4 M NH_4OH, 2/2/1, (4) CH_2Cl_2/methanol, 1/1; (5) acetone/ethyl acetate, 9/1. Similarly, a glyoxal standard is coincident with a less abundant ^{14}C-labelled product in the same five solvent systems. Except for some amino acids the balance of the labelled material has not been positively identified.

Total soluble organic products from several experiments were concentrated and fractionated by chromatography on 8 × 150 mm AG 50 × 8 H^+ ion exchange columns. The water rinse (four-bed volumes) released 90% of the input counts as the neutral and basic fractions. A final rinse of four bed volumes of 1 M NH_4OH released 9% of the input counts (total released 17,620 d.p.m. ^{14}C), the acidic fraction; while 1% of the input label remained bound to the column. The acidic fraction was concentrated and separated by two-dimensional silica gel-H TLC in (I) methanol/chloroform/17% NH_4OH, 2/2/1, and then (II) *n*-butanol/acetic acid/water, 8/2/2. The plate was stained with ninhydrin to detect visually amino acid spots. Then the plate was divided into 22 equal area zones, each of which was eluted and measured for ^{14}C activity. Seven well-defined ninhydrin-positive spots were observed with ^{14}C activities of 2,600, 1,200, 660, 510, 460, 200 d.p.m. of the total 17,620 d.p.m. placed on the plate. The balance of the labelled ninhydrin-negative material migrated with an $R_f > 0.9$ in both solvent systems. The most intense ninhydrin positive spots which bore ^{14}C label corresponded to: glycine (2,600 d.p.m.), alanine

Table 1 UV photochemical reduction of carbonate by model protocells

Experiment no.	Total input ^{14}C label (mCi)	Total recovered ^{14}C organic (mCi $\times 10^{-6}$)	Organic C in product (mM $\times 10^{-9}$)	No. C atoms reduced (erg^{-1} cm^{-2} $\times 10^3$)	Φ_{app} : apparent quantum yield
1	0.5	0.36	7.2	12	0.09×10^{-6}
2	0.5	0.045	0.9	—	—
3	0.4	92.66	1,853.2	3,096	24.2×10^{-6}
4	0.125	3.36	67.2	112	0.88×10^{-6}
5	0.5	20.35	407.0	680	5.32×10^{-6}

Experimental procedures are detailed in text. Experiment 1, no organic microstructures added; experiment 2, no UV irradiation; experiment 3, all reactants present (microstructures in distilled water, $^{14}CO_3^{2-}$, N_2/H_2 head space); experiment 4, all reactants present but microstructures disrupted by repeated freeze–thaw cycles; experiment 5, all reactants present but *p*H 8.1 50 mM phosphate buffer replaced distilled water. The freeze–thaw cycling of structures before irradiation disrupted mechanically the phase boundaries of these model protocells.

(1,200 d.p.m.), β-alanine (1,100 d.p.m.) and valine (660 d.p.m.). Thus ~3.3% of the total ^{14}C-labelled soluble products appear to be amino acids. About four other ninhydrin-negative ^{14}C-labelled spots (450, 2,500, 1,850, 1,500 d.p.m.) remain unidentified.

The source of nitrogen for these amino acids is undetermined and could be derived either from microstructure organic material or from atmospheric N_2.

Freshly made organic microstructures washed twice in H_2O were pressed dry and 9 mg were placed in a quartz reaction vessel with 1 mg of unlabelled $NaHCO_3$. The vessel was flushed with N_2 and filled with 710 mm N_2, 50 mm H_2 and 100 µl ^3H_2O (100 mCi at a specific activity of 1 Ci ml^{-1}, Research Products). The mixture was stirred and irradiated for 24 h as in previous experiments. After irradiation, the reaction mixture was lyophilized and resuspended in H_2O for three cycles to remove unreacted tritium label. Then 2 ml of H_2O containing 56,000 d.p.m. ^{14}C-labelled formaldehyde (specific activity 30 Ci M^{-1}) were added as an internal standard. $Ba(OH)_2$ in excess was added to precipitate carbonate, the supernatant was lyophilized to 50 µl and applied to two silica gel-H plates for development in CCl_2H_2/methanol solvent systems at 95/1 and 1/1 (v/v). Zones 2 cm wide were eluted and measured for both ^{14}C and ^3H activity by liquid scintillation counting. Total organic ^3H recovered from the reaction was 1.113×10^6 d.p.m. Of this, >70% was found to co-elute in both TLC solvent systems with the internal standard formaldehyde. The tritium label experiment shows that 1.67×10^{-5} mM ^3H was found in formaldehyde, while a typical carbon label experiment (as experiment 3, Table 1) yields 1.85×10^{-6} mM carbon in formaldehyde. As the experimental recovery of formaldehyde is low and variable, these two experiments indicate only qualitatively that many formaldehyde protons appear to be derived from water.

Routine microbiological surveys for contamination, conducted by placing aliquots of organic microstructure suspensions directly onto agar media (trypticase soy agar, nutrient agar, yeast extract agar) were negative, as were quartz tube reaction mixtures. Unirradiated reaction mixtures containing ^{14}C-labelled $NaHCO_3$ yielded only traces of presumptive labelled organic products (Table 1, experiment 2, 100 d.p.m. recovered). UV-irradiated reaction mixtures which contained ^{14}C label but no microstructures yielded some presumptive organic material (Table 1, experiment (1), 800 d.p.m. recovered), but this value is orders of magnitude less than in experiments which include microstructures.

Hong, Hong and Becker[10] had previously shown that amino acids and other biomolecules can be formed from single carbon precursors by photochemically generated 'hot' hydrogen atoms. Sagan and Khare[11] have also demonstrated amino acid synthesis by the long-wavelength UV irradiation of two carbon atom precursors photosensitized by H_2S. Both reports indicate quantum yields of the order of 10^{-5}. The apparent quantum yields we report here ($\sim2.4\times10^{-5}$) are based on incident UV 254-nm photons and are uncorrected for absorption, reflection and shielding and are not directly comparable with those of Sagan *et al.*[11] or of Hong *et al.*[10]. Although no specific photosensitizers, such as H_2S or Hg, were present in our system, organic microstructures themselves absorb UV and visible light effectively (0.05 mg dry wt. has an $A_{254_{nm}}$ of ~2.5). Presumably it is on the surfaces of these structures where photosensitization, perhaps to form free radicals[12], occurs. Tyson and Oyama[13] demonstrated incorporation of gaseous $^{14}CO_2$ into UV photolysis products of solid amino acids at wavelengths >220 nm with apparent quantum yields of about 4×10^{-6} to 4×10^{-7}. Apparently aliphatic amino acids form free radicals when irradiated at 254 nm, which might then react with CO_2 (ref. 13). Hubbard, Hardy and Horowitz[15] and Tseng and Chang[12] showed the production of organic compounds, principally formaldehyde and glycolic acid, on irradiation over a broad spectral range below 300 nm in surface catalysed reactions of CO, CO_2 and H_2O.

This work has been suppórted by NASA research grant NGR 12-001-109 and by the University of Hawaii.

Received 5 February; accepted 10 April 1981.

1. Folsome, C. E. *The Origin of Life* (Freeman, San Francisco, 1979).
2. Schopf, J. W. *A. Rev. Earth planet. Sci.* **3**, 213 (1975).
3. Herrara, A. L. *Arch. Plasmologie Gen.* **1**, 55 (1912).
4. Oparin, A. I. *Origin of Life* (Dover, New York, 1953).
5. Bahadur, K. *Agra Univ. J. Res. (Sci.)* **14**, 41 (1965).
6. Fox, S. W. *Naturwissenschaften* **60**, 359 (1973).
7. Yanigawa, H. & Egami, F. in *Origin of Life* (ed. Noda, H.) 385 (Centre for Academic Publishing, Japan, 1978).
8. Folsome, C. E. *Naturwissenschaften* **63**, 303 (1976).
9. McLaren, A. D. & Shugar, D. *Photochemistry of Proteins and Nucleic Acids* (Pergamon, New York, 1964).
10. Hong, K., Hong, J. & Becker, R. S. *Science* **184**, 985 (1974).
11. Sagan, C. & Khare, B. N. *Science* **173**, 417 (1971).
12. Tseng, S. S. & Chang, S. *Nature* **248**, 575 (1974).
13. Tyson, B. J. & Oyama, V. I. *Curr. mod. Biol.* **5**, 98 (1973).
14. Johns, R. B. *Biochem. biophys. Acta* **147**, 369 (1967).
15. Hubbard, J. S., Hardy, J. B. & Horowitz, N. H. *Proc. natn. Acad. Sci. U.S.A.* **68**, 574 (1971).

Structurally preserved angiosperm flowers from the Upper Cretaceous of southern Sweden

E. M. Friis* & A. Skarby†

*Botany Department, Bedford College, University of London, Regent's Park, London NW1 4NS, UK
†Geologiska Institutionen, Kungstensgatan 45, Box 6801, 11 386 Stockholm, Sweden

The angiosperms, or flowering plants, are believed to have originated and diversified during the Cretaceous. The earliest angiosperm fossils are pollen and leaf impressions found in late early Cretaceous strata[1,2]. We now describe the discovery of fossil flowers in the Upper Cretaceous of Sweden which represent a significant addition to our very limited knowledge of the floral structure of Cretaceous angiosperms.

Among morphological and stratigraphic studies of the early angiosperm record[1,3–6], the investigations by Doyle and Hickey[1,3] on the pollen and leaves from the Potomac Group of eastern North America have provided critical evidence on the early diversification of the group. Angiosperm reproductive structures in the Cretaceous are mainly fruits and seeds, and floral structures are extremely rare[7–9], with the result that work on evolutionary relationships within the angiosperms[10,11] has largely concentrated on living plants. Most Cretaceous flowers so far described are impression or compression fossils[9] which usually yield only incomplete information of floral structures and arrangement of parts. Stopes and Fujii[12] described a petrified gynoecium from the late Cretaceous of Japan, but the perianth parts of the flowers were incompletely preserved and the organization of this structure, attributed to the Liliaceae, is not clear. The only other structurally preserved flower known to be from the Cretaceous[13] was found in North America. This is preserved as a three-dimensional compaction and rather detailed information on the floral structure has been obtained, although the degree of compression limits the extent to which its organization can be reconstructed.

Our discovery of angiosperm flowers from Sweden represents the best-preserved Cretaceous floral structures available. The flowers were sieved from kaolin-rich sediments collected in Höganäs AB's quarry at Åsen, north-east Scania, Sweden, together with other fossil plant remains. They are usually found in cross-bedded, poorly sorted sands and their orientation along the sedimentary structures indicates transportation together with the mineral grains. The deposits are dated on palynological evidence as Upper Santonian or Lower Campanian; the flowers must be of about the same age. The sediments are considered to be of fluviatile origin deposited close to the Senonian shoreline[14,15]. The fossil flowers are a minor part of the organic remains obtained from the sieving, the commonest of the fossils

0028–0836/81/240484—03$01.00

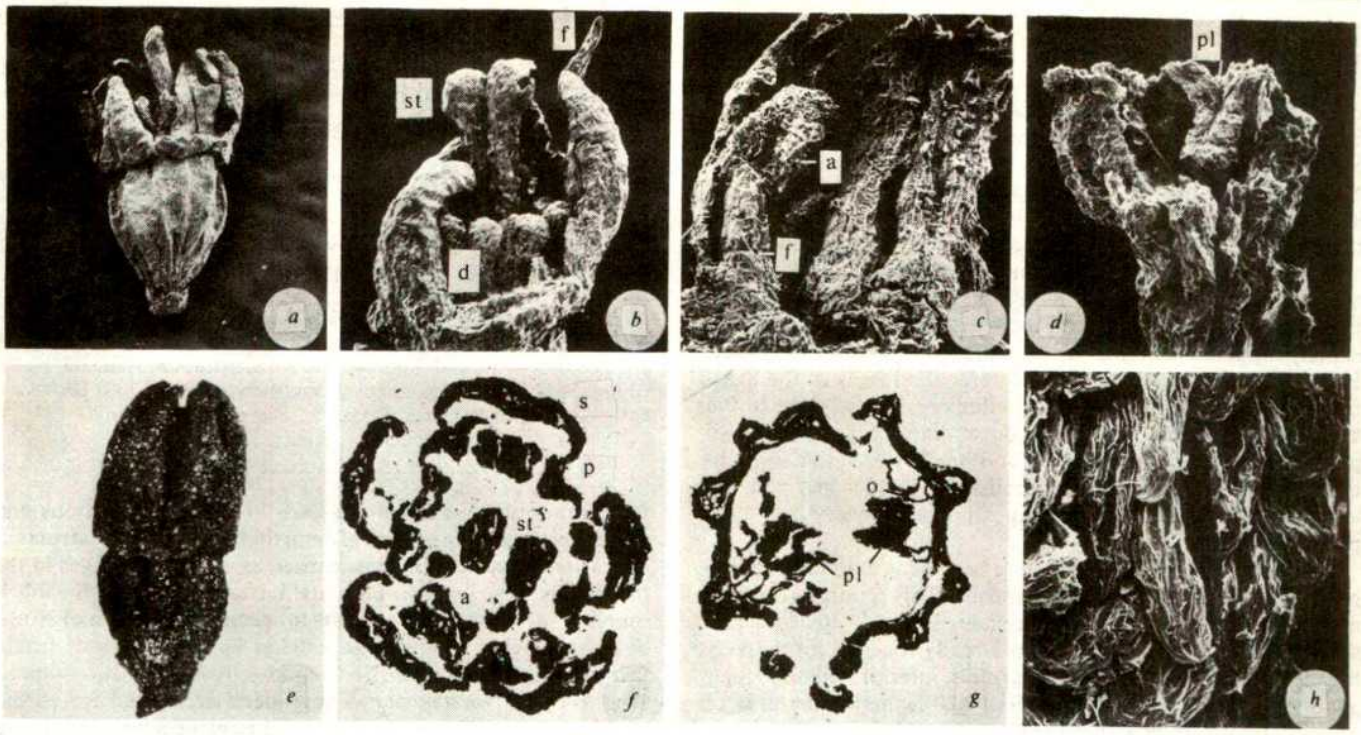


Fig. 1 Fossil flowers from the Upper Cretaceous of Scania, southern Sweden. *a*, Scanning electron micrograph of mature flower with one sepal missing, showing fragment of petal and filament (×21). *b*, Fragment of flower showing prominent disk (d) and styles with capitate stigma (st); the apex of one filament (f) projects beyond the sepals (×60). *c*, Part of flower bud with perianth removed, showing stamen with filament (f) and anther (a) preserved (×110). *d*, Fragment of ovary showing apical and lateral placenta (pl) (×55). *e*, Flower bud showing sepals and inferior ovary (×28). *f*, Transection of flower bud shown in *e* with whorls of sepals (s), petals (p), anthers (a) and styles (st) (×60). *g*, Transection through inferior ovary showing two placentae (pl) and ovules (o) (×60). *h*, Ovules attached to the placenta, removed from the ovary (×230). The specimens will be deposited at the Department of Geology, University of Stockholm, Sweden.

being fruits and seeds. Conifer leaves of unknown affinity are common in some samples, and megaspores, sporangia and dispersed anthers have also been recovered, although the angiosperm remains are the dominant element. Palynological studies[16,17] show that the angiosperms are also predominant among the palynomorphs, and pollen forms referred to the Normapolles complex are common.

One flower type represented by several developmental stages from small flower buds to mature flowers (Fig. 1*a*, *e*) has been examined in detail using the scanning electron microscope, as well as by serial sectioning for light microscopy. The flowers are preserved in a three-dimensional state, which we attribute to their charcoal composition. Thin sections of the tissue show homogenization of the cell wall, a feature which is considered to be characteristic for fusinite (charcoal)[18]. Charring probably occurred within the litter at the soil surface as a result of forest fire, before transport and incorporation in the sediment. Such transformation to charcoal evidently confers rigidity and resistance to collapse on burial and several records of relatively delicate plant parts preserved three-dimensionally in charcoal are now known[19,20].

The fossil flowers are ~ 2 mm long and 1 mm broad. From the degree of opening of the perianth, most of the flowers must have been close to maturity, even though so minute. All floral parts except the anthers seem to be persistent so that examination of the androecium was restricted mainly to the flower buds. The flowers are regular and bisexual with radial symmetry (actinomorphic) and with the floral parts in whorls. The sepals seem to be more resistant than the petals and are more frequently preserved. Both sepals and petals number five and are free (Fig. 1*f*), the calyx being imbricate and the corolla of open aestivation. The androecium is apparently diplostemonous and composed of 10 stamens. The anthers are dorsi-fixed and introrse with longitudinal dehiscence (Fig. 1*c*). The pollen grains are slightly prolate, tricolpate, and very small (~10 μm in equatorial diameter), with a tectate exine. They are probably not mature.

The gynoecium is inferior with a unilocular ovary formed by two fused carpels. This inferior ovary is characterized by 10 longitudinal ribs (Fig. 1*a*, *e*). The gynoecium is surmounted by a prominent 10-lobed disk (Fig. 1*b*) which is thought to be a nectary; because of this, and on account of the floral structure generally, the flowers are believed to have been pollinated by insects. The ovary has two free styles, each terminated by a small capitate stigma (Fig. 1*b*). The fruit is a capsule splitting apically between the carpels and there are two apical and lateral, pendant placentae each with many ovules (Fig. 1*d*, *g*, *h*). The

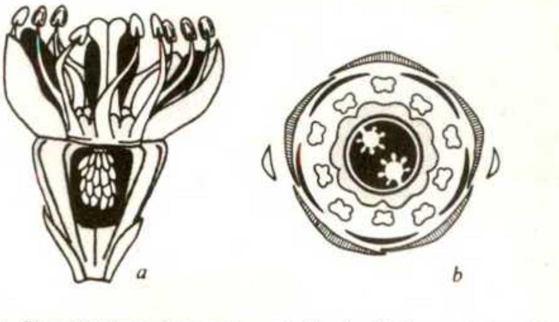


Fig. 2 Simplified reconstruction of the fossil flower (*a*), with part of ovary wall removed to show one placenta and floral diagram (*b*).

seeds are small and thin-walled with slightly elongated surface cells. The arrangement of the floral parts is shown in Fig. 2*a*, *b*.

The organization and structure of the fossil flower suggest a close relationship to a number of Recent genera of the Saxifragales. Diplostemonous flowers are known within the Saxifragales, although obdiplostemony or haplostemony is the common organization of the androecium. Unilocular and inferior ovaries are also found in a number of saxifragalean families. Free, pendant placentae, however, are not common, and have only been observed in a few genera. Although all characters known to occur in the fossil flower are found within the saxifragalean complex, no modern genus seems to possess exactly the same combination of characters as the fossil flower. However, it seems appropriate to refer the fossils to the Saxifragales and it is believed that the flowers may be close to the source of several modern saxifragalean lines. Among the fossil plant remains from our material another flower with affinity to the Saxifragales has been discovered. It is distinguished from the above described flower in having a bilocular ovary and axile placentation.

The fossil flowers described here represent the oldest epigynous flower known in the fossil record. This report also documents the presence of the Saxifragales in the Cretaceous and demonstrates that the group had already reached a relatively high level of evolution, with syncarpous, inferior and unilocular gynoecia. Although the attainment of the major features of angiosperm diversity within the Cretaceous is suggested by leaf and pollen characters[1,3], this is the first demonstration of these particular advanced floral features within the Cretaceous.

The Saxifragales are represented in the Tertiary by a number of genera and by the end of the Eocene the presence of four families, the Iteaceae, Pittosporaceae, Hydrangeaceae and Saxifragaceae, have been documented by the occurrence of floral structures from the Baltic Amber[21]. Our discovery of saxifragalean flowers in the Cretaceous thus demonstrates that some of the diversity of the group seen in the Lower Tertiary was already evident in the Cretaceous. This is consistent with the central position of the Saxifragales in angiosperm phylogeny postulated by Takhtajan[10] and Hutchinson[11].

We thank Professor W. G. Chaloner of Bedford College, University of London, for helpful suggestions and critical discussion, Professor K. Larsen, University of Aarhus, for helpful discussion, and Höganäs AB for help and material. E.M.F. acknowledges the support of the Danish Natural Science Research Council and the receipt of a British Council Research fellowship.

Received 15 January; accepted 26 March 1981.

1. Doyle, J. A. & Hickey, L. J. in *Origin and Early Evolution of Angiosperms* (ed. Beck, C. B.) 139–206 (Columbia University Press, New York, 1976).
2. Hughes, N. F. *Bot. Rev.* **43,** 105–127 (1977).
3. Hickey, L. J. & Doyle, J. A. *Bot. Rev.* **43,** 3–104 (1977).
4. Beck, C. B. (ed.) *Origin and Early Evolution of Angiosperms,* 1–341 (Columbia University Press, New York, 1976).
5. Kemp, E. M. *Palaeontology* **11,** 421–434 (1968).
6. Wolfe, J. A., Doyle, J. A. & Page, V. M. *Ann. Mo. bot. Gdn* **62,** 801–824 (1975).
7. Knobloch, E. *Sb. geol. Ved, Rada TG* **19,** 79–132 (1977).
8. Dilcher, D. L., Crepet, W. L., Beeker, C. D. & Reynolds, H. C. *Science* **191,** 845–856 (1976).
9. Dilcher, D. L. *Rev. Palaeobot. Palynol.* **27,** 291–328 (1979).
10. Takhtajan, A. L. *Flowering Plants, Origin and Dispersal,* 1–310 (Oliver & Boyd, Edinburgh, 1969).
11. Hutchinson, J. *Evolution and Phylogeny of Flowering Plants,* 1–717 (Academic, London and New York, 1969).
12. Stopes, M. C. & Fujii, K. *Phil. Trans. R. Soc.* **B201,** 1–90 (1910).
13. Tiffney, B. H. *Nature* **265,** 136–137 (1977).
14. Grönwall, K. A. *Sver. geol. Unders. Afh.* C**261,** 1–185 (1915).
15. Lundegren, A. *Geol. För. Stockh. Förh.* **53,** 298–320 (1931).
16. Ross, N. E. *Bull. geol. Instn Univ. Upsala* **34,** 25–43 (1949).
17. Skarby, A. *Acta Univ. Stockh.* **16,** 1–60 (1968).
18. Cope, M. J. & Chaloner, W. G. *Nature* **283,** 647–649 (1980).
19. Scott, A. C. & Collinson, M. E. in *Scanning Electron Microscopy in the Study of Sediments* (ed. Whally, B.) 137–167 (Geo-Abstracts, Norwich, 1978).
20. Alvin, K. L. *Palaeontology* **17,** 587–598 (1974).
21. Conwentz, H. *Die Flora des Bernsteins, 2. Die Angiospermen des Bersteins,* 1–140 (Danzig, 1886).

0028–0836/81/240486—04$01.00

Congruent shifts in sand eel abundance in western and eastern North Atlantic ecosystems

K. Sherman*, C. Jones*, L. Sullivan*, W. Smith†, P. Berrien† & L. Ejsymont‡

* National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Narragansett, Rhode Island 02882, USA
† National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Sandy Hook, New Jersey 07732, USA
‡ Morski Instytut Rybacki, Zaklad Sortowania i Oznaczania Planktonu, Szczecin, Poland

It has been suggested, on the basis of model simulations and slowly accumulating empirical data, that changes in the structure of marine ecosystems may be caused as much by changes in the trophic levels as by environmental factors[1,2]. Support for this is found in recently observed shifts in species abundance of North Sea fish stocks, where large catches in the 1970s of small, fast-growing, opportunistic plankton-feeding fishes—sprat, sand eel and Norway pout—have been accounted for as the result of the replacement by these species of depleted herring and mackerel stocks[3,4]. This suggestion is difficult to confirm, as knowledge of the abundance of these 'opportunistic' species since the mid-1960s is largely limited to catch data for the North Sea[5]. The only sand eel report independent of commercial fisheries catch data is for larvae collected in 1948–68, just before the sharp decline in mackerel and herring stocks[6]. The area surveyed did not, however, cover all the major population centres in the North Sea. In support of replacement, we describe here evidence from the north-west Atlantic which indicates that population explosions of small, fast-growing sand eel can coincide with depletions of larger tertiary predators, including herring and mackerel in a continental shelf ecosystem. Sand eels serve as an important link in marine food chains, preying on secondary producers and constituting an important food of higher trophic level fish and marine mammals[7,8].

Our results are based on 6 years (1974–79) of ichthyoplankton surveys, conducted by the United States, Poland and the Soviet Union on the continental shelf off the east coast of the United States, as part of a joint MARMAP study of shelf productivity. [MARMAP is the first multinational programme for measuring long-term changes in key structural components of very large marine ecosystems. At present 260,000 km^2 of the north-west Atlantic continental shelf are being sampled routinely to measure changes in key structural components, including primary production (^{14}C), chlorophyll *a* phaeophytin, nutrients (NO_2, NO_3, SiO_3, NH_4 and PO_4), zooplankton, ichthyoplankton, fish, benthos, seabirds, water-column temperature, salinity and circulation[9].] Between 6 and 12 surveys were done per year. The sampling area covered the continental shelf and included four sub-areas: Gulf of Maine, Georges Bank, southern New England and middle Atlantic Bight (see Fig. 1). Ichthyoplankton were collected at about 180 locations situated 25–35 km apart and during the 6 yr, we sampled ~6.5 million km^2.

The data presented here are based on population densities of sand eel calculated for the annual spawning period in winter, when sand eel larvae, *Ammodytes* spp., are most abundant. Two species, *A. dubius* and *A. americanus*, reportedly occur in the survey area[10]. Their taxonomy remains confused and we are uncertain whether our collections included only one or both. Thus, until the identity of the larvae is clarified, they are classified as *Ammodytes*.

At each sampling location, tows for ichthyoplankton and zooplankton were made through the water column using a paired bongo-type plankton sampler[11] with 60-cm openings and

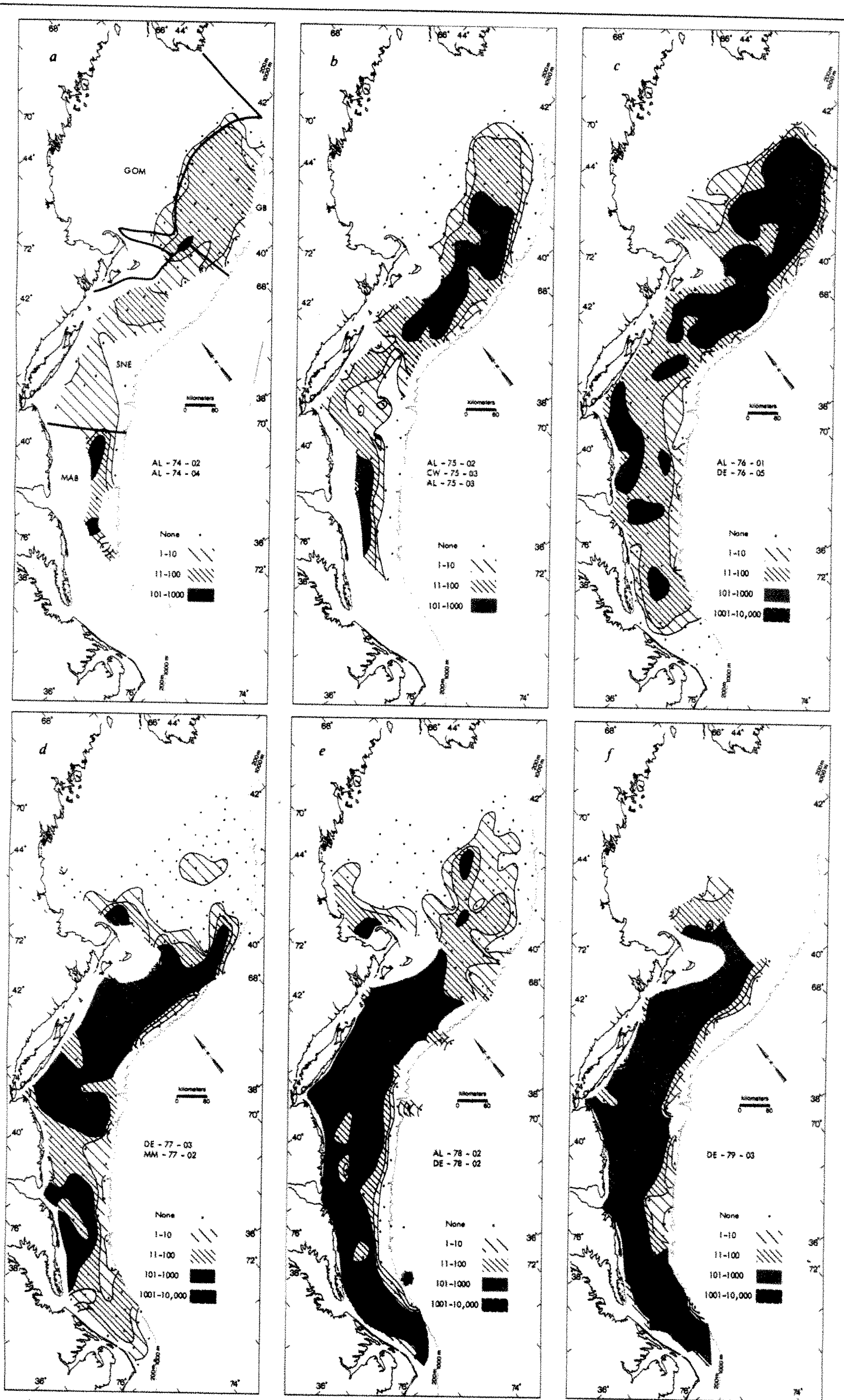


Fig. 1 Increase in abundance of sand eel larvae (*Ammodytes*) on the continental shelf off the north-east United States coast in winter 1974–1979. The four sub-areas of the shelf sampled include the Gulf of Maine (GOM), Georges Bank (GB), southern New England (SNE) and mid-Atlantic Bight (MAB). Samples were collected from the NOAA research vessels: Albatross IV (AL), Delaware II (DE), Commonwealth (CW) and Mount Mitchell (MM). *a*, 11 February–22 March 1974; *b*, 12 February–16 March 1975; *c*, 10 February–24 March 1976; *d*, 12 February–8 April 1977; *e*, 14 February–17 March 1978; *f*, 23 February–15 March 1979. Shading represents number of *Ammodytes* larvae per 10 m^2 surface area.

Table 1 Winter abundance estimates of sand launce, *Ammodytes*, larvae in three sub-areas off northeastern United States

Year		Georges Bank ($41,809\ km^2$)	Southern New England ($59,906\ km^2$)	Middle Atlantic ($58,326\ km^2$)	Total ($\times 10^{10}$)
1974	k	33.941	11.216	48.692	
	s.e.	1.054	1.093	1.726	
	Abundance	14.190	6.719	28.400	49.309
	% Of larval fish	55	13	88	
1975	k	90.427	51.555	57.044	
	s.e.	1.602	1.220	1.371	
	Abundance	37.806	30.884	33.271	101.961
	% Of larval fish	59	60	90	
1976	k	1,018.311	243.798	55.393	
	s.e.	1.990	1.440	1.297	
	Abundance	425.746	146.050	32.309	604.105
	% Of larval fish	97	96	84	
1977	k	9.931	603.255	76.991	
	s.e.	1.996	1.446	1.404	
	Abundance	4.152	361.386	44.905	410.443
	% Of larval fish	88	81	71	
1978	k	13.916	569.359	313.297	
	s.e.	1.122	1.451	1.332	
	Abundance	5.818	341.080	182.733	529.631
	% Of larval fish	91	92	94	
1979	k	442.324	738.239	577.487	
	s.e.	1.934	1.402	1.561	
	Abundance	184.931*	442.250	336.825	964.006
	% Of larval fish	98	79	87	

Abundance estimates were determined by Δ-distribution[25]. Retransformed mean abundance (k) is expressed as larvae per $10\ m^2$ surface area. Larval abundance ($\times 10^{10}$) is expansion of k to reflect the sub-area size. (See Berrien *et al.*[26] for discussion of rationale and procedures for use of Δ-distribution.)

* Estimate based on limited sampling.

nets of 0.333 and 0.505 mm mesh. These nets were towed at ship speeds ranging from 1.5 to 3.5 knots. The bongo sampler was lowered obliquely from the surface to within 5 m of the bottom at a wire speed of $50\ m\ min^{-1}$ and retrieved at $20\ m\ min^{-1}$. Water filtered through the net was measured with a flow meter and a time–depth recorder was used to measure the towing path of the sampler. The ichthyoplankton samples were sorted, species identified and enumerated, and measured at the Plankton Sorting Center, Szczecin, Poland. The data on ichthyoplankton were calculated initially as numbers of larvae per $10\ m^2$ of sea surface, and expanded to numbers of larvae per km^2 and numbers of larvae $\times 10^{10}$ for the entire shelf.

The north-west Atlantic ecosystem off the northeastern United States has been subjected to significant fishing stress; fish biomass in the region was reduced by 50% from 1968 to 1975 (ref. 12). Since then, the silver hake and squid stocks have been increasing but herring and mackerel stocks remain low[13]. From 1974 to 1979, coincident with the low mackerel and herring stocks, we observed a population explosion of sand eel larvae increasing from a low shelf density of 49×10^{10} in 1974 to 964×10^{10} in 1979. During this 6-yr period the percentage of sand eel increased from less than 50% of the total mid-winter ichthyoplankton community to more than 85%. A summary of comparisons of abundance levels and percentage composition of sand eel in the ichthyoplankton community among areas and in years is made in Table 1.

The centre of sand eel larval abundance shifted during the years of study. Abundance was low in 1974 and 1975. In 1976 the centre of population density was on Georges Bank. Abundance estimates were significantly lower on Georges Bank in

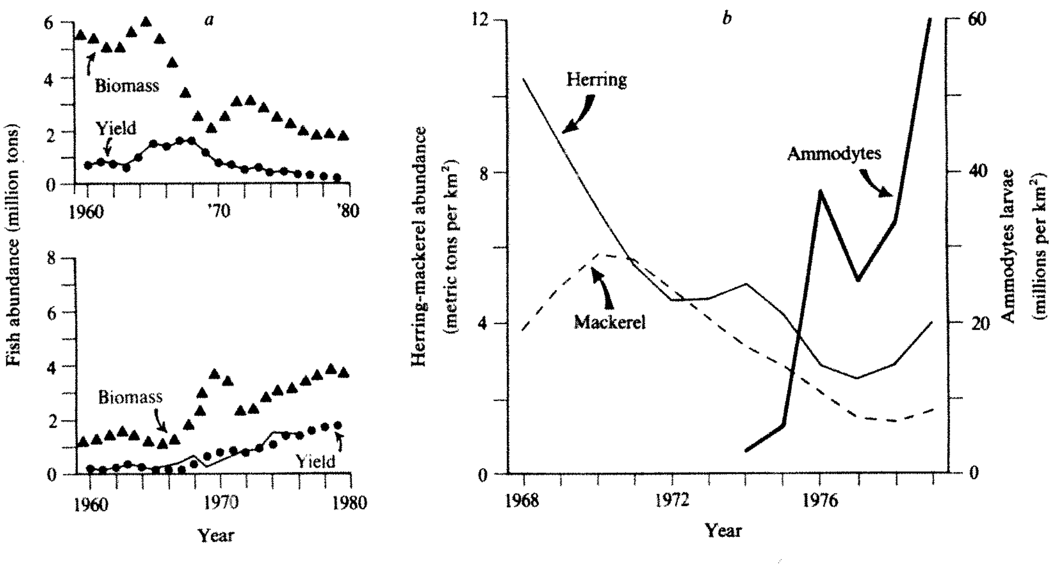


Fig. 2 *a*, Decline in herring and mackerel biomass (▲) and yield (●) (upper graphs), and successive rise in biomass (▲) and yield (●) (lower graphs) of small, fast-growing sprat, sand eels and Norway pout in the North Sea from 1960 to 1979 as depicted by Ursin[3]. *b*, Decline in herring (———) and mackerel (- - -) biomass in metric tons per km^2 and successive increase in larval *Ammodytes* (———) abundance in millions of larvae per km^2 for the north-west Atlantic shelf from Cape Hatteras to the Gulf of Maine (1974–1979). Herring densities represent stocks ranging from southern New England to western Nova Scotia and are based on an updating (Waring, personal communication) of the data of Sissenwine and Waring[27]; mackerel densities include stocks ranging from Cape Hatteras to Labrador (1968–1979), based on the data of Anderson[28].

1977 and 1978 but larvea of other taxa were also greatly reduced in number and the percentage contribution of sand eels remained close to 90%. Because of technical difficulties we were unable to complete our early winter survey in 1979. After 1976, the principal centres of abundance were in the southern New England and mid-Atlantic Bight areas; large patches of larvae in densities of over 1,000 per 10 m^2 extending for distances greater than 300 km occurred most frequently off the southern New England coast (Fig. 1).

In the North Sea, the fishery yield for juvenile and adult sand eel, sprat and Norway pout increased from <200,000 metric tons in 1960 to 1.6 million metric tons in 1976. In the absence of a fishery in the north-west Atlantic, it is difficult to estimate the contribution of larval abundance of sand eel to the biomass of adults. Based on available evidence, we conclude that the adult stock is increasing. (1) MARMAP bottom trawl surveys have shown an increasing trend in abundance in southern New England and the mid-Atlantic Bight[14]. (2) An increase in abundance has been observed directly by divers[15], and (3) a significant increase has been observed in the consumption of sand eel by predators from 1968 to 1976 (ref. 16).

The influence of temperature has been suggested as a principal factor in controlling abundance levels of fish stocks on the continental shelves of the north-west Atlantic off the Canadian maritimes and the northeastern United States[17,18]. The reductions in herring and mackerel on both sides of the Atlantic in response to heavy fishing mortality, followed by increases in sand eel and other small, fast-growing fish as shown in Fig. 2*a*, *b* make unlikely the hypothesis that the changes are due to environmental factors. Our findings show that when a large biomass of mid-size predators is removed it can be replaced by smaller, faster-growing, opportunistic species as described for the North Sea[3,4] and observed by us.

The full consequence of structural change with respect to earlier unperturbed status has not yet been documented for large marine ecosystems. We are just starting to obtain the multitrophic level data needed to unscramble inter- and intraspecific ecosystem relationships[9,19-22]. The influence of structural changes on aquatic ecosystems can result in significant economic and aesthetic loss. In a species succession model for the Great Lakes, removal of large predators resulted in the proliferation of fast-growing smaller alewives[23]. A size-component response to structural modification in a large marine ecosystem has been observed in the Antarctic where, based on recent population trends[24], the decrease in the biomass of large blue whales may be in the process of replacement by the faster-growing smaller minke whales and seals. Monitoring of the shelf ecosystem in the north-west Atlantic is continuing in an attempt to measure the impact of any increase in the demersal and pelagic fish-predator field (for example, cod, haddock, hake, mackerel and herring) on the changing structure of the shelf ecosystem.

Received 29 December 1980; accepted 17 March 1981.

1. Steele, J. H. & Frost, B. W. *Phil. Trans. R. Soc.* **B280,** 485–534 (1977).
2. Daan, N. *Rapp. P.-v. Réun. Cons. perm. int. Explor. Mer* **177,** (in the press).
3. Ursin, E. *Cons. int. Explor. Mer C. M. F.* **42** (1977).
4. Andersen, K. P. & Ursin, E. *Rapp. P.-v. Réun. Cons. perm. int. Explor. Mer* **172,** 286–291 (1978).
5. Hempel, G. *Rapp. P.-v. Réun. Cons. perm. int. Explor. Mer* **172,** 445–449 (1978).
6. Hart, P. J. B. in *The Early Life History of Fish* (ed. Blaxter, J. H. S.) 171–182 (Springer, Berlin, 1974).
7. Bigelow, H. B. & Schroeder, W. C. *Fish. Bull. U.S.A.* **74,** 577 (1954).
8. Reay, P. J. *F. A. O. Fish. Biol. Synopses* **82** (1970).
9. Sherman, K. in *Advanced Concepts on Ocean Measurements for Marine Biology* (eds Diemer, F. P., Vernberg, F. J. & Mirkes, D. Z.) 9–37 (Belle W. Baruch Institute for Marine Biology and Coastal Research, Univ. of S. C. Press, 1980).
10. Richards, S. W., Perlmutter, A. & McAneny, D. C. *Copeia 1963,* 358–377 (1963).
11. Posgay, J. A. & Marak, R. R. *J. Northw. Atl. Fish. Sci.* **1,** 91–99 (1980).
12. Clark, S. H. & Brown, B. E. *Fish. Bull. U.S.A.* **75,** 1–21 (1977).
13. Clark, S. H. & Brown, B. E. *Investigación pesq.* **43,** 107–122 (1979).
14. Grosslein, M., Langton, R. W. & Sissenwine, M. P. *Rapp. P.-v. Réun. Cons. perm. int. Explor. Mer* **177,** (in the press).
15. Meyer, T. L., Cooper, R. A. & Langton, R. W. *Fish. Bull. U.S.A.* **77,** 243–253 (1979).
16. Bowman, R. *NMFS/NEFC Woods Hole Lab. Ref.* No. 80–05 (1980).
17. Dow, R. L. *J. Cons. int. Explor. Mer* **37,** 274–280 (1977).
18. Sutcliffe, W. H. J., Drinkwater, K. & Muir, B. S. *J. Fish. Res. Bd Can.* **34,** 19–30 (1977).
19. Anderson, K. P. & Ursin, E. *Medd. Dan. Fish.-Havunder. N.S.* **7,** 319–435 (1977).
20. Regier, H. A. & Henderson, H. F. *Trans. Am. Fish. Soc.* **102,** 56–72 (1973).
21. Mills, E. L. & Fournier, R. O. *Mar. Biol.* **54,** 101–108 (1979).
22. Cohen, E., Grosslein, M., Sissenwine, M. & Steime, F. *Cons. int. Explor. Mer C.M.L.* **64** (1980).
23. Smith, S. H. *Bull. Fish. Res. Bd Can.* **25,** 667–693 (1968).
24. Beddington, J. R. & Grenfell, B. T. in *The Management of the Southern Ocean* (eds Mitchell, B. & Sandbrook, R.) 31–40 (International Institute for Environment and Development, London, 1980).
25. Atchinson, J. *J. Am. statist. Ass.* **50,** 272–901 (1955).
26. Berrien, P. L., Naplin, N. A. & Pennington, M. R. in *The Early Life History of Fish II* (eds Lasker, R. & Sherman, K.); *Rapp. P.-v. Réun. Cons. perm. int. Explor. Mer* **178** (in the press).
27. Sissenwine, M. P. & Waring, L. L. *NMFS/NEFC Woods Hole Lab. Ref.* No. 79–12 (1979).
28. Anderson, E. D. *NMFS/NEFC Woods Hole Lab. Ref.* No. 80–29 (1980).

Monocyclic β-lactam antibiotics produced by bacteria

R. B. Sykes, C. M. Cimarusti, D. P. Bonner, K. Bush, D. M. Floyd, N. H. Georgopapadakou, W. H. Koster, W. C. Liu, W. L. Parker, P. A. Principe, M. L. Rathnum, W. A. Slusarchyk, W. H. Trejo & J. S. Wells

The Squibb Institute for Medical Research, PO Box 4000,
Princeton, New Jersey 08540, USA

A recent article[1] described two novel monocyclic β-lactam antibiotics produced by bacteria. In the past few years, we have screened bacteria isolated from numerous ecosystems and habitats for their ability to produce β-lactam antibiotics, using a strain of *Bacillus licheniformis* which is specific for molecules containing a β-lactam and which is sensitive down to 100 ng ml^{-1}. From over one million bacterial isolates screened, we have now identified seven related molecules containing β-lactams produced by a range of bacterial species. For this new family of monocyclic β-lactams we suggest the class name 'monobactam'. Bacteria which produce β-lactams are, in reality, relatively common in nature.

The β-lactam most frequently detected in our studies was SQ 26,445 (Fig. 1, I), one of the compounds isolated from a strain of *Pseudomonas acidophila*[1]. β-Lactam-producing strains isolated from 41 separate locations were identified as species of

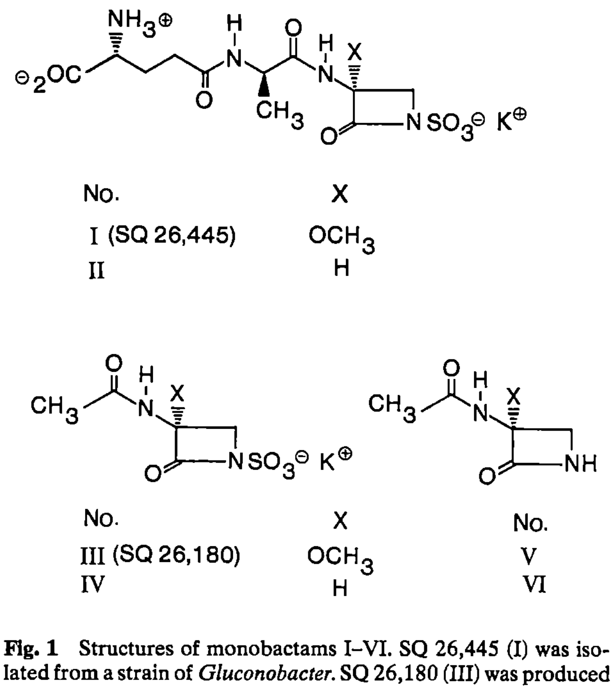


Fig. 1 Structures of monobactams I–VI. SQ 26,445 (I) was isolated from a strain of *Gluconobacter*. SQ 26,180 (III) was produced by strains of *Chromobacterium violaceum*.

Table 1 Antibacterial activity of monobactams

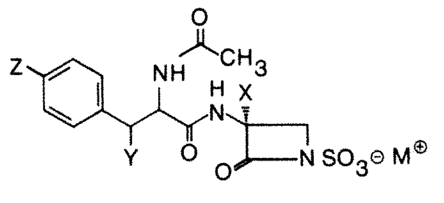
		MIC (μg ml^{-1})		
Organism	SC	I	III	IV
Staphylococcus aureus	1,276	>100	50	50
Escherichia coli	10,896	25	25	>100
Klebsiella pneumoniae	9,527	50	>100	>100
Proteus mirabilis	3,855	12.5	>100	>100
Enterobacter cloacae	8,236	25	>100	>100
Pseudomonas aeruginosa	9,545	50	3.1	>100

Minimum inhibitory concentrations (MICs) were determined by a twofold agar dilution method on DST agar (Oxoid). Final inoculum level was 10^4 colony-forming units.

Gluconobacter or *Acetobacter*, and the original producing strain of SQ 26,445 as a polar-flagellated, oxidative Gram-negative rod. This culture was cytochrome oxidase negative, positive for catalase, citrate and esculin hydrolysis and accumulated poly-β-hydroxybutyric acid. Growth was good at *p*H 4.5 in malt yeast extract medium; acid production was observed as a zone of clearing around colonies grown on yeast extract/glucose/calcium carbonate agar. The organism was assigned to the genus *Gluconobacter* on the basis of these characteristics.

For isolation of SQ 26,445 fermentations were carried out at 25 °C for 18–24 h with aeration and agitation in a 380-l fermenter containing 250 l of medium (0.5% yeast extract and 1.0% glucose in distilled water). SQ 26,445 was purified by a combination of ion-exchange, charcoal, and reverse-phase chromatography. The structure (Fig. 1, I) was assigned by amino acid analysis of hydrolysis products (6 M HCl), N-terminal analysis[2] and NMR spectroscopic comparison to γ-Glu-Ala[3] and SQ 26,180 (*vide infra*).

The simplest monobactam found was SQ 26,180 (Fig. 1, III) produced by strains of *Chromobacterium violaceum*. Although many thousands of *C. violaceum* strains were examined, organisms producing III were isolated from only eight locations; each of these was identified as *C. violaceum* on the basis of the following characteristics. The organism was an aerobic Gram-negative rod, motile due to a single polar flagellum with an occasional single smaller lateral flagellum. Colonies of this bacterium gave a violet pigment on nutrient agar and production of pigment was enhanced by the addition of either tryptophan or yeast extract. Good growth was observed between 15 and 37 °C, with no growth below 4 °C or above 37 °C. Casein was hydrolysed strongly and esculin hydrolysis was negative. Glucose, fructose and trehalose were fermented; L-arabinose was not used fermentatively or oxidatively; HCN was produced.



No.	X	Y	Z	M
VII	OCH_3	H	H	Na
VIII	OCH_3	H	OH	K
IX	H	H	OH	K
X	OCH_3	OH	$OSO_3^{\ominus}Na^{\oplus}$	Na
XI	OCH_3	$OSO_3^{\ominus}Na^{\oplus}$	$OSO_3^{\ominus}Na^{\oplus}$	Na

Fig. 2 Structures of monobactams VII–XI, produced by *Agrobacterium radiobacter*.

For isolation of SQ 26,180, fermentations were carried out at 25 °C for 18–24 h with aeration and agitation in a 380-l fermenter containing 250 l of medium (yeast extract, 0.25%; glucose, 3.0%; N-Z amine-A, 1.0% in tap water). SQ 26,180 was concentrated from broth by ion-pair extraction and purified by partition, ion-exchange, size-exclusion and reverse-phase chromatography. The structure (Fig. 1, III) was assigned by spectroscopic deduction and confirmed by synthesis from azetidinone V (refs 4, 5) which also demonstrated the absolute configuration. Thus, treatment of V with dimethylformamide-SO_3 complex, followed by workup with buffered tetra-*n*-butyl ammonium bisulphate, gave the tetra-*n*-butyl ammonium salt of III, which was identical to material prepared from crystalline SQ 26,180, melting point 194 °C (dec). The unmethoxylated analogue IV was similarly prepared from azetidinone VI (ref. 5).

Table 2 Susceptibility to and inhibition of β-lactamases

Compound	Enzyme	Relative V_{max}	K_m (mM)	I_{50} (mM)
Benzylpenicillin	*S. aureus* (PCl)	100	0.03	—
III		<0.01	—	>2.0
IV		0.71	0.04	—
VIII		<0.02	—	>0.4
IX		12	0.98	—
Cephaloridine	*E. coli* (TEM-2)	100	0.66	—
III		<0.01	—	>2.0
IV		12	0.68	—
VIII		<0.01	—	>0.4
IX		10	0.28	—
Cephaloridine	*Klebsiella* (K-1)	100	0.15	—
III		<0.04	—	>2.0
IV		86	0.45	—
VIII		<0.03	—	>0.4
IX		220	1.3	—
Cephaloridine	*Enterobacter* (P-99)	100	0.58	—
III		0.06	—	0.22
IV		0.46	0.41	—
VIII		<0.01	—	0.002
IX		0.15	0.19	—

All studies were performed at 25 °C in 0.1 M phosphate buffer, *p*H 7.0, using homogeneous TEM-2, P-99 and *S. aureus* β-lactamases. K-1 β-lactamase was a partially purified preparation containing a single β-lactamase activity. Monobactams were prepared in 20 mM KH_2PO_4 and diluted in 0.1 M phosphate buffer before assays. β-Lactamase activities were determined using a Gilford 250 spectrophotometer at the following wavelengths: 295 nm for cephaloridine; 240 nm for benzylpenicillin; 206 nm for compounds III, IV and VIII and 235 nm for IX. Kinetic constants were obtained from Lineweaver–Burk plots of at least four concentrations for each substrate. Relative V_{max} values were calculated by setting V_{max} for cephaloridine or benzylpenicillin to 100. I_{50} values, the concentration of compound required to inhibit enzyme activity by 50%, were determined using 1.0 mM cephaloridine as the substrate for P-99 β-lactamase and 0.5 mM benzylpenicillin as the substrate for TEM-2, K-1 and *S. aureus* β-lactamases.

Agrobacterium radiobacter SC 11,742 produces a mixture of monocyclic β-lactams, five of which have been identified as VII–XI (Fig. 2). Compound IX is the first naturally occurring non-methoxylated derivative isolated in quantity.

The β-lactam-producing organism was identified as *A. radiobacter* on the basis of the following. It was an aerobic Gram-negative rod, motile by means of sub-polar and peritrichous flagella. The strain was cytochrome oxidase positive, arginine dihydrolase negative and produced 3-ketolactose from lactose[6]. Growth of the organism on mannitol yeast extract agar containing Congo red resulted in dye uptake[7]. Fluorescence was not observed, and there was no pathogenicity for abraded stems of tomato and sunflower seedlings.

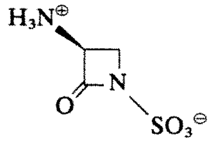
For isolation of compounds VII–XI, fermentations were carried out at 25 °C for 40–45 h with aeration and agitation in a 380-l fermenter containing 250 l of medium yeast extract (0.5% and glucose, 1.0% in distilled water). The mixture of VII–XI was isolated from broth by ion-pair extraction. Chromatography on Sephadex G-10 in water gave rough separation according to charge type (XI eluted first, followed by X and finally a mixture of VII–IX). Further purification by combinations of ion-exchange and HP-20 chromatography gave pure samples of VII–XI. Structures were assigned to these by elemental composition and spectroscopic properties. The structures of IX and II were confirmed by synthesis from 6-aminopenicillanic acid (6-APA) (to be described elsewhere). Our assignments of C-3 absolute configuration in I, VII, VIII, X and XI rest on the analogy with III; side-chain absolute configuration has only been rigorously determined for I and IX (D-Tyr). Compound IX is thought to be the precursor of VIII (ref. 8); indeed, we have detected small quantities of II and IV by HPLC in the fermentation broths from which we isolated I and III.

Antimicrobial activities of selected monobactams are shown in Table 1. Although the compounds are weakly active against a range of bacteria, the methoxylated compound III shows greater activity against Gram-negative organisms, especially *Pseudomonas*, compared with the non-methoxylated derivative IV. Compounds VII–XI showed similar antimicrobial properties to those observed for compounds III and IV.

The methoxylated monobactams III and VIII showed a high degree of stability to β-lactamases from both Gram-positive and Gram-negative organisms, which agrees with previous work[1], whereas the desmethoxy monobactams IV and IX were readily hydrolysed by enzymes with primarily penicillinase activities. The *Enterobacter* P-99 cephalosporinase was the only enzyme in this series to show significant affinity for all the monobactams tested (Table 2).

The intrinsic activity of III and IV was assessed on the basis of binding to essential penicillin binding proteins (PBPs) in *Escherichia coli*. The PBP profile indicated a lack of affinity (up to 100 μg ml^{-1}) for any of the essential PBPs. Interestingly, methoxylation of the monobactam nucleus increased binding to the non-essential PBP 4 and 5/6 (D, D-carboxypeptidase IB and IA, respectively)[9,10], an effect similar to that observed for cephalosporins.

On the basis of the novel nucleus XII (3-aminomonobactamic acid, 3-AMA) we have synthesized a highly active β-lactamase-stable derivative which has been developed for clinical evaluation and will be reported elsewhere.



XII (3-AMA)

Received 5 March; accepted 23 March 1981.

1. Imada, A. *et al. Nature* **289,** 590–591 (1981).
2. Schroeder, W. A. & Le Gette, J. *J. Am. chem. Soc.* **75,** 4612–4615 (1953).
3. Kristensen, I. & Larsen, P. O. *Acta chem. scand.* **27,** 3123–3125 (1973).
4. Gordon, E. M. *et al. J. Am. chem. Soc.* **102,** 1690–1702 (1980).
5. Kamiya, T. *Recent Advances in the Chemistry of β-Lactam Antibiotics* (ed. Elks, J.) 281–298 (Chemical Society, London, 1977).
6. Bernaerts, M. J. & DeLey, J. *Nature* **197,** 406–407 (1963).
7. Hahn, N. J. *Can. J. Microbiol.* **12,** 725–733 (1966).
8. O'Sullivan, J. *et al. Biochem. J.* **179,** 47–52 (1979).
9. Iwaya, M. & Strominger, J. L. *Proc. natn. Acad. Sci. U.S.A.* **74,** 2980–2984 (1977).
10. Spratt, B. G. & Strominger, J. L. *J. Bact.* **127,** 660–663 (1976).

Vasopressor receptor antagonist prevents behavioural effects of vasopressin

M. Le Moal*, G. F. Koob*, L. Y. Koda*, F. E. Bloom*, M. Manning†, W. H. Sawyer‡ & J. Rivier§

*Arthur Vining Davis Center for Behavioral Neurobiology and the
§Peptide Biology Laboratory, The Salk Institute, PO Box 85800, San Diego, California 92138, USA
†Department of Biochemistry, Medical College of Ohio, Toledo, Ohio 43699, USA
‡Department of Pharmacology, College of Physicians, Columbia University, New York, New York 10032, USA

The antidiuretic hormone vasopressin[1] also has a direct vasoconstrictor effect on the smooth muscles of the vascular system[2]. This may be physiologically significant[3], but higher doses of the hormone are required in conscious subjects than when the autonomic nervous system is depressed by anaesthesia or ganglionic blocking agents[2]. Vasopressin may also function as a releaser of corticotropin[4,5], in the maintenance of learned avoidance behaviours and in memory consolidation[6–8]. Although these functions are controversial[9–11], it is of interest that corticotropin[12,13] and other peptides cleaved from the propiocortin[14] precursor also influence avoidance behaviours[13]. Here, we have examined a highly potent peptide analogue[15] of arginine vasopressin (AVP) for its ability to antagonize the pressor and behavioural responses to AVP in conscious rats. Our results indicate that doses of AVP which prolong extinction of active avoidance behaviour for several hours after subcutaneous injection[8,13,16], also produce early pressor responses. Conversely, we also find that doses of the antagonist peptide, 1-deaminopenicillamine, 2-(*O*-methyl)tyrosine AVP (dPTyr(Me)AVP), which can prevent the pressor response[15], also abolish the effects of subcutaneously injected AVP on prolongation of extinction. This blockade indicates that signals from peripheral visceral sources may have an important role in the subsequent behavioural changes, and suggests that the receptors at which AVP elicits its pressor effect are similar to those leading to its behavioural action.

For measurements on blood pressure in conscious animals, Wistar rats (Charles River Farms, 300–400 g) were anaesthetized with halothane through an open face mask. The abdominal aorta was cannulated with PE 10 polyethylene tubing through the left femoral artery, and tied in place about 2.5 cm above the entrance to the abdominal cavity. The tubing, filled with heparinized saline (20 units ml^{-1}), was then plugged distally and led through the subcutaneous space to exit through a dorsal incision at the inter-scapular level. Cannulae were flushed daily with heparinized saline and remained patent for at least 5 days. Beginning 24 h or more after cannulation, each rat was placed in a sound-attenuated experimental chamber and the cannula connected to a Statham pressure transducer (P2310) and monitored with a Buxco (Quincy) cardiovascular analyser for heart rate and systolic, diastolic and mean arterial pressure. Each rat was used only once a day. Blood pressure was monitored at various times after subcutaneous injection of saline or various doses of AVP or dPTyr(Me)AVP.

Argine vasopressin was synthesized at the Salk Institute (J.R.). The vasopressin analogue dPTyr(Me)AVP was synthesized first at the Medical College of Ohio (M.M., W.H.S.) and subsequently at the Salk Institute (J.R.). This compound has been fully characterized as a specific antagonist to the pressor actions of vasopressin in the urethane-anaesthetized rat[15]; in

0028-0836/81/240491—03$01.00

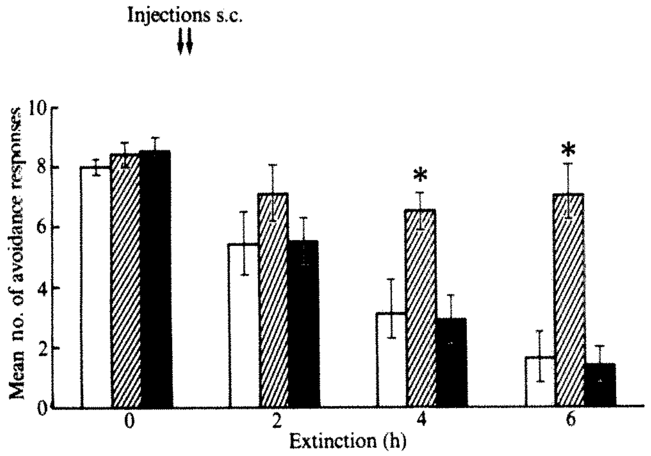


Fig. 1 Effects of AVP and AVP plus dPTyr(Me)AVP on extinction of active avoidance behaviour. After 3 days of training, rats meeting the criteria of at least seven successful avoidances during the first 10 trials of extinction were injected subcutaneously (s.c.) with saline (twice at 2.0-min intervals), AVP plus saline or AVP plus dPTyr(Me)AVP immediately after the first set of 10 extinction trials on day 4. Rats receiving AVP only show persistent avoidance throughout the 6 h of observation, when tested on 10 trials at each of the three next 2-h intervals, whereas rats receiving either saline with no peptide or both peptides extinguished this active avoidance behaviour at similar rates. □, Saline+saline ($n = 9$); ▨, saline+AVP, 6 μg per kg ($n = 9$); ■, dPTyr(Me)AVP, 30 μg per kg+AVP, 6 μg per kg. *Significantly different from both the saline with no peptide and both peptides groups; $P < 0.05$ Newman–Keuls test following analysis of variance.

this preparation AVP is a more potent pressor agonist than in unanaesthetized rats[2]. Before injection, peptides were first dissolved in 50 μl of 0.01 M hydrochloric acid and then further diluted in 0.9% NaCl.

For behavioural experiments rats (140–160 g) were handled for 30 s for 2–5 days after arrival in the laboratory. They were trained and tested in a modified Skinner box (BRS/LVE) 25 cm × 30 cm × 27 cm containing a grid floor, composed of 16 bars, 0.5 cm in diameter and spaced 2 cm apart, for administering shock. A wooden pole, blackened with ink and 1.67 cm in diameter, extended 25 cm from the cage floor to the ceiling; a 50-W incandescent light bulb was centred just above the clear Plexiglass ceiling and served as the conditioned stimulus.

For acquisition, each rat was subjected to 10 trials per day for 3 days. The inter-trial interval was 40, 60 or 80 s, with an average of 60 s. A trial began with illumination of the conditioned stimulus followed 5 s later by a scrambled a.c. shock (0.3–0.7 mA) delivered to the grid floor, lasting up to 30 s or until the rat jumped on the pole. During the first three trials on the first day of acquisition, rats failing to jump on the pole within 5–10 s were gently placed on the pole by the experimenter. Throughout training, rats failing to climb down the pole within 30 s after jumping were gently turned around on the pole and thus forced to return to the grid floor. All testing was accomplished in a darkened room between 1000 and 1900 h. For extinction, the same procedure was used except that the shock was omitted and the rats were subjected to 10 trials every 2 h during the fourth day. Rats were selected for drug injections by a similar procedure to that reported by De Wied and associates[8]. Only rats that made seven or more avoidance responses during the first 10 extinction trials were treated with peptide or placebo immediately after the session and continued in the experiment through extinction trials. All injections were subcutaneous, in the neck, with a constant volume of 0.5 ml of peptide solution. Control rats received saline alone. All behavioural testing was performed according to a 'blind' protocol in which the experimenter testing the rats was unaware of the subject's treatment.

Results on systolic blood pressure or on extinction, for a given time point, were analysed with a single-factor analysis of variance (ANOVA). Subsequent individual mean comparisons were made using the Newman–Keuls test, but only at those time points showing a significant ANOVA value, that is, $P < 0.05$.

In confirmation of previously published observations by the De Wied group[8,12,13] and our own[16], subcutaneous injections of as little as 1 μg AVP (~6 μg per kg) gave reproducible prolongation of the extinction of active (pole jump) avoidance which was highly significant for as long as 6 h (see ref. 16 and Fig. 1). We therefore examined this dosage range for effects on arterial pressure in the conscious rat (Figs 2, 3), and found that both 5 and 25 μg per kg doses of AVP led to prompt elevations of arterial blood pressure without overt behavioural effects. These effects peaked at 10–20 min after injection and were generally within normal ranges after 1 h (data not shown). Using the threshold pressor dose of AVP at 5 μg per kg, we evaluated various doses of the antagonist peptide and found that the pressor response was significantly depressed and delayed with 5 μg per kg dPTyr(Me)AVP, and apparently completely blocked by a 25 μg per kg dose (Fig. 3). A 10-fold higher dose of the antagonist also produced a significant (but equally transient) lowering of basal systolic pressure (Fig. 2). As the 25 μg per kg dose of antagonist completely blocked the response to a 5 μg per kg dose of AVP but did not itself cause hypotension, the 5:1 ratio of antagonist to agonist was used in further tests of AVP in active avoidance extinction (Fig. 1); here, the co-injection of dPTyr(Me)AVP with AVP completely prevented the prolongation of extinction at 4 or 6 h after the injection. Rats receiving AVP and the antagonist therefore did not differ from saline-injected rats and showed normal extinction behaviour, whereas rats receiving AVP and saline showed markedly prolonged extinction, throughout the 6 h of testing.

Although the mechanism by which AVP and several other peptides alter extinction behaviour in active or passive avoidance paradigms is unknown, it has been hypothesized that such actions are centrally elicited because intracerebroventricular injections of even smaller amounts of these peptides produce these same behavioural effects. However, even with intracerebroventricular injections it is impossible to know whether

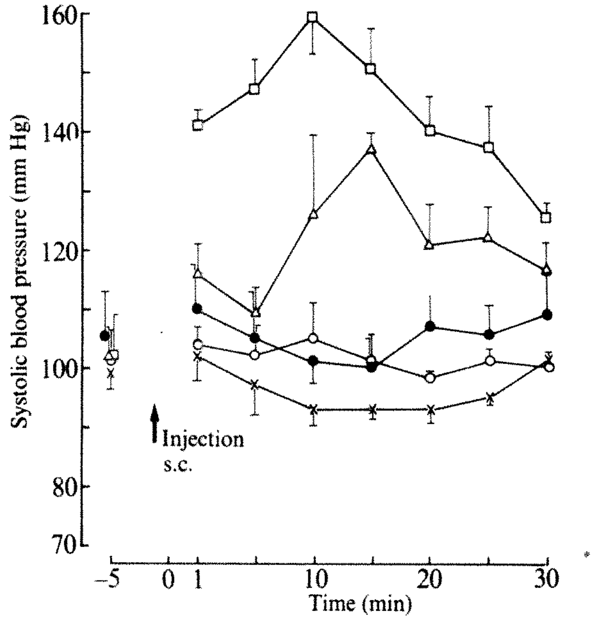


Fig. 2 Effects of AVP on systolic blood pressure in conscious rats, as recorded by direct pressure transduction through chronically implanted arterial cannulae before and after the subcutaneous injections indicated. After AVP doses of 5 or 25 μg per kg, blood pressures were rapidly and significantly elevated; 250 μg per kg doses of the antagonist peptide significantly lowered basal arterial pressure. □, AVP, 25 μg per kg ($n = 3$); △, AVP, 5 μg per kg ($n = 3$); ●, AVP, 1 μg per kg ($n = 3$); ×, dPTyr(Me)AVP, 250 μg per kg ($n = 3$); ○, saline ($n = 3$).

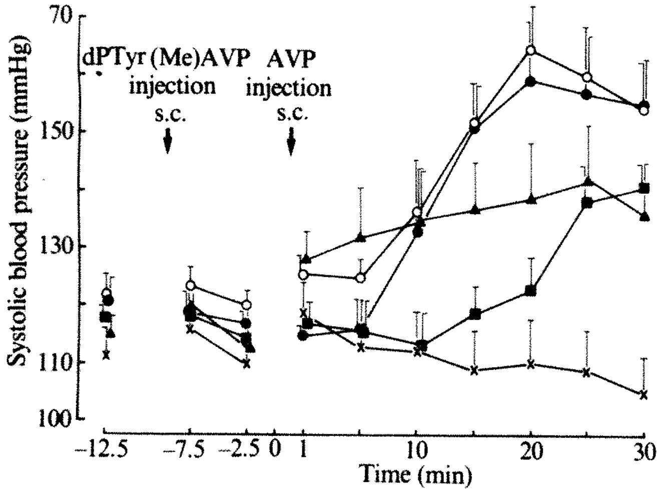


Fig. 3 Effects of dPTyr(Me)AVP on increases of systolic blood pressure produced by AVP. The antagonist peptide at various doses was injected subcutaneously 7.5 min before the subcutaneous injection of a constant dose (5 μg per kg) of AVP. Compared with saline-injected controls ('0 μg per kg' antagonist), blood pressure responses were significantly depressed by 1 and 5 μg per kg doses of antagonist and completely antagonized by the 25 μg per kg dose. Doses of dPTyr(Me)AVP (μg per kg): ○, 0 ($n=3$); ●, 0.3 ($n=3$); ▲, 1.0 ($n=4$); ■, 5.0 ($n=6$); ×, 25 ($n=4$).

SCIENCE COLLEGE
Upper Circular Road
CALCUTTA

the targets responsible for the altered behaviour are within the substance of the brain or within the ventricular spaces. Examples of the latter sort of target, which would be open to contact with circulating blood or extracellular space, are the circumventricular organ systems[17]. Although blood–brain barriers are reduced in some brain regions, such as hypothalamus and area postrema, these areas are not generally regarded as critical for memory processes. The present results provide additional clues to the possible sites and mechanisms of the peptide-induced behavioural alterations, as they indicate that in the conscious rat, doses of AVP injected peripherally elicit a significant rise in arterial pressure, and that when this pressor response is prevented by an analogue of AVP, the behavioural effects of AVP are also antagonized.

Our data could be interpreted to indicate that the vascular response, and perhaps other visceral effects, including evoked release of corticotropin[4,5], could lead to altered visceral afferent signals which are sufficient to alert the rat and to cause the persistence of ongoing or escape behaviours. For example, an acute induced hypertensive state can reduce reactivity to noxious stimuli through a baroreceptor-mediated reduction of cerebral arousal[18]. Under the latter view, the effects of AVP may be more generally interpreted as a hidden signal within a holistic set of integrative central nervous system visceral afferent information. Under extremes of dehydration or acute vascular hypovolaemia, such behavioural effects of AVP could be an important adjunct to hormone actions directed to restoration of fluid volume and general homeostasis[3]. Alternatively, the blockade of the behavioural response following blockade of the pressor response could simply indicate that AVP acts both centrally and peripherally and that the two events, pressor response and prolongation of extinction, are unrelated. In this view, even if the peripheral vascular effects of AVP are not directly the cause of the subsequent behavioural effects, it remains possible that the integrative actions of endogenous AVP as a physiological regulator elicit a state of central activation appropriate for stressful situations which are generated in parallel with the peripheral pressor action. Additional observations are therefore required to determine whether AVP and other peripherally injected peptides alter behaviour by directly activating central operational mechanisms such as memory or retrieval or whether they induce altered behaviours by triggering peripheral receptors which in turn evoke some altered perception of visceral status that is sufficient to place the animal in a state of alerted surveillance.

We thank N. Callahan for manuscript preparation. Studies were supported by USPHS grants DA 01785 and AA 03504.

Received 5 February; accepted 16 March 1981.

1. Kamm, O., Aldrich, T. B., Grote, I. W., Rowe, L. W. & Bugbee, E. P. *J. Am. chem. Soc.* **50,** 573–601 (1928).
2. Hays, R. M. in *The Pharmacological Basis of Therapeutics* (eds Gilman, A. G. & Goodman, L. S.) 916–928 (Macmillan, New York, 1980).
3. Strauss, M. C. *Body Water in Man* (Little Brown, Boston, 1957).
4. Pearlmutter, A. F., Rapino, E. & Saffran, M. *Neuroendocrinology* **15,** 106–119 (1974).
5. Gillies, G. E. & Lowry, P. J. *Nature* **278,** 463–464 (1979).
6. Ader, R. & De Wied, D. *Psychon. Sci.* **29,** 46–48 (1972).
7. Bohus, B., Kovacs, G. L. & De Wied, D. *Neuropharmacology* **17,** 239–247 (1978).
8. Walter, R., Van Ree, J. M. & De Wied, D. *Proc. natn. Acad. Sci. U.S.A.* **75,** 2493–2496 (1978).
9. Bailey, W. H. & Weiss, J. M. *Brain Res.* **162,** 174–178 (1979).
10. Hostetter, G., Jubb, S. L. & Kozlowski, G. R. *Neuroendocrinology* **30,** 174–177 (1980).
11. Schmidt, W. K., Holaday, J. W., Loh, H. H. & Way, E. L. *Life Sci.* **23,** 151–158 (1978).
12. De Wied, D. *Proc. Soc. exp. Biol. Med.* **122,** 28–32 (1966).
13. De Wied, D. *Life Sci.* **20,** 195–204 (1977).
14. Rubinstein, M., Stein, S. & Udenfriend, S. *Proc. natn. Acad. Sci. U.S.A.* **75,** 669–673 (1978).
15. Bankoski, K., Manning, M., Haldar, J. & Sawyer, W. H. *J. med. Chem.* **21,** 850–853 (1978).
16. Koob, G. F. *et al. Regulatory Peptides* (in the press).
17. Weindl, A. in *Frontiers in Neuroendocrinology* (eds Ganong, W. F. & Martin, L.) 3–32 (Oxford, New York, 1973).
18. Dworkin, B. R., Fylewich, R. J., Miller, N. E., Craigmyle, N. & Pickering, T. G. *Science* **205,** 1299–1301 (1980).

Chronic morphine administration increases the apparent number of α_2-adrenergic receptors in rat brain

Margaret Hamburg & John F. Tallman

Section on Biochemistry and Pharmacology, Biological Psychiatry Branch, National Institute of Mental Health, 9000 Rockville Pike, Bethesda, Maryland 20205, USA

When the α_2-adrenergic agent[1] clonidine (Catapresan) is given to opiate addicts undergoing withdrawal, a rapid and dramatic reversal of objective signs and subjective symptoms of withdrawal can be observed[2], which include a marked decrease in hypertension and anxiety levels with no accompanying euphoria[3]. Such clinical findings have been correlated with noradrenergic neuronal activity in the locus coeruleus, and systemic or local administration of clonidine inhibits the firing of locus coeruleus noradrenergic cells[4]; administration of opiates produces a similar inhibiting effect. Stimulation of the locus coeruleus either electrically or pharmacologically (by the administration of α-adrenergic or opiate antagonists) increases the firing rate of these neurones[4] and elicits many of the symptoms of opiate withdrawal[5]. Both morphine and clonidine have been shown to block the effects of electrical and pharmacological activation of the locus coeruleus in primates and rats[4]. However, single-unit recording techniques indicate that opiates and clonidine elicit their action at independent binding sites[5–7], although they may activate a common mechanism. We have explored the possible relationship between opiates and clonidine at the receptor level, and demonstrate here that chronic, but not acute, administration of morphine to rats results in increased binding of clonidine to α_2-adrenergic binding sites in brain. Increased binding of clonidine is due to an enhanced number of α_2-adrenergic binding sites but no change in their affinity.

Male Sprague–Dawley rats (200–250 g) under light halothane anaesthesia were implanted subcutaneously with pellets containing 75 mg morphine sulphate (or control pellets). On the second day, two additional pellets were implanted[8], and after

Table 1 Effect of chronic and acute morphine treatment on adrenergic binding

	Control	Treated	%Change
	specifically bound (fmol per incubation ± s.e.m.)		
Chronic morphine			
Cortex			
^{3}H-clonidine	69.7 ± 3.0	93.2 ± 3.7*	+33.7
^{3}H-WB-4101	107.9±3.4	102.7 ± 1.9†	−4.8
^{3}H-Dihydroalprenolol	29.2±0.7	34.4 ± 1.9*	+17.9
Brain stem			
^{3}H-Clonidine	32.4±2.1	48.2 ± 1.0*	+48.4
^{3}H-Dihydroalprenolol	22.8±1.7	29.5 ± 0.8‡	+29.5
Acute morphine			
Cortex			
^{3}H-Clonidine	73.8±3.9	70.5 ± 3.7†	−4
^{3}H-Dihydroalprenolol	32.6±1.3	35.1 ± 1.6†	+7

Brains were weighed then homogenized in 40 vol (w/v) of ice-cold 0.05 M Tris-HCl buffer (*p*H 7.4 at 4°C). The homogenate was centrifuged at 18,000*g* for 20 min, the supernatant decanted and the pellet rehomogenized. After two such washes, the pellet was resuspended in 10 vol (original) of buffer and used in the assay. 1-[^{3}H] dihydroalprenolol (β-adrenergic receptors) and ^{3}H-WB-4101 (α-adrenergic receptors) binding were assayed as described earlier[10–12] except that 1 nM tritiated ligand was used. ^{3}H-clonidine binding (α_2-adrenergic receptors) was determined by a modification of a published method[10,11]. For the case of clonidine binding, 20 mg tissue weight equivalents of homogenate (200 μl) were used. Incubations were carried out at room temperature in 0.05 M Tris-HCl (*p*H 7.4 at 25°C) containing 2.0 nM ^{3}H-clonidine (NEN, and in early experiments the gift of Boehringer Ingelheim) in a total volume of 10.0 ml. Nonspecific binding is defined as that binding occurring in the presence of 10^{-5} M phentolamine. Incubations were for 20 min duration and were terminated by rapid filtration through Whatman GFB filters followed by two washes (6 ml each) with Tris buffer. Filters were dried out and counted in a liquid scintillation counter using an appropriate counting fluor after several hours in the dark. In the conditions described above, all three incubations were linear with protein concentratrations at all ligand concentrations used here and in Fig. 1 and in the sites behaved in pharmacological displacement experiments as described in the literature[10,11]. In particular, the ^{3}H-clonidine site showed the following order of displacement (at both 0.5 nM and 5 nM concentration): clonidine > adrenaline > noradrenaline > piperoxone > yohimbine > phenylephrine > WB-4104 > prazosin. This site also demonstrates modulation by GTP and analogues, some mono- and divalent cations, but not ATP (confirming the findings of refs 10, 11). Dissociation of bound ^{3}H-clonidine is biphasic in the absence of GTP, indicating the possible presence of multiple forms of α_2-adrenergic receptors (see also Fig. 1 legend); in the presence of GTP (5×10^{-4} M), however, dissociation of bound clonidine was rapid, complete and monophasic. Conditions (multiple washes, large incubations volumes) were chosen to minimize these confounding variables in the study of the morphine effects; their relevance to the functioning of α_2-adrenergic receptors in intact cells may be great but it is not the subject of this report. Additionally, in more concentrated incubations with incompletely washed fractions, they may contribute to nonlinearity in Scatchard plots or added variability in the assay. Data are presented as fmol per incubation and the relative results are the same if calculated per mg protein or per g wet weight of tissue. ^{3}H-Clonidine binding is about 15% higher than reported values due to the different incubation volume.

$^{*}P < 0.05$, $n = 5$; † not significantly different from controls; ‡$P < 0.01$ $n = 5$.

two additional days (72 h total), opiate withdrawal was tested in a few animals of each group either by the injection of naloxone (2 mg per kg) or removal of the pellet[9]. Successful treatment was indicated by the appearance of such signs as wet dog shakes, jumping, writhing and diarrhoea (1 h after naloxone or 24 h after pellet removal). The remainder of the animals in each group were killed, their brains immediately removed, dissected and frozen on dry ice.

This chronic morphine treatment (Table 1) produced an increase in ^{3}H-clonidine binding (α_2-adrenergic receptors[10,11]) to brain membranes prepared from cerebral cortex and brain stem (Table 1). As previously reported and shown here in a typical experiment (other replications are mentioned below and in Fig. 1), a smaller (but statistically significant) enhancement of 1-[^{3}H]dihydroalprenolol (DHA) binding (β-adrenergic receptors[12]) was also found[13], although no significant alteration in ^{3}H-WB-4101 binding[10-12] (α_1-adrenergic receptors) was observed. Following acute administration of morphine sulphate (50 mg per kg, intraperitoneally, 1 h previously), no significant changes in either clonidine or DHA binding were observed. Additionally, there were only small increases in clonidine binding after 1 day. The full effect takes several days to appear, consistent with the reported changes in noradrenaline turnover and cell firing[2-4]. Analyses of data obtained in studies of clonidine binding to membranes prepared from chronic morphine-treated animals and controls indicated that the enhanced binding of clonidine was due to an apparent increase in the number of binding sites (Fig. 1). This was also the case (data not shown) with DHA, as demonstrated previously[13]. A single injection of morphine had no effect on either clonidine or DHA binding (Table 1), and the inclusion of up to 10^{-5} M morphine in the assay itself had no effect on control binding of either drug (data not shown).

Following morphine treatment, a single dose of naloxone will elicit severe withdrawal symptoms in the conditions described above. Additionally, within 1 h, naloxone administration entirely reversed the enhancement of both clonidine and DHA binding obtained after chronic morphine treatment (in this experiment, a 29.5 ± 0.5% enhancement of clonidine binding ($n = 5$) and a 16.5 ± 0.3% enhancement of DHA binding were obtained after morphine treatment; $P < 0.05$ in each case). Following naloxone, morphine-treated animals ($n = 5$) also treated with naloxone showed a 8 ± 1% increase in clonidine binding which was not different from that in animals treated with naloxone alone (7 ± 1% increase compared with untreated animals ($n = 5$)). DHA binding was increased 4 ± 1% following naloxone and 7 ± 1% in morphine-treated animals also treated with naloxone. The addition of naloxone (10^{-6} M) to binding assays with tissue from either control or morphine-treated animals did not lower binding of clonidine or DHA, suggesting that naloxone had a more complex action than direct occupancy of opiate receptors leading immediately to altered clonidine binding. Interestingly, pretreatment of animals with naloxone had little effect on the binding of clonidine and DHA when compared with unpretreated control animals. For pharmacok-

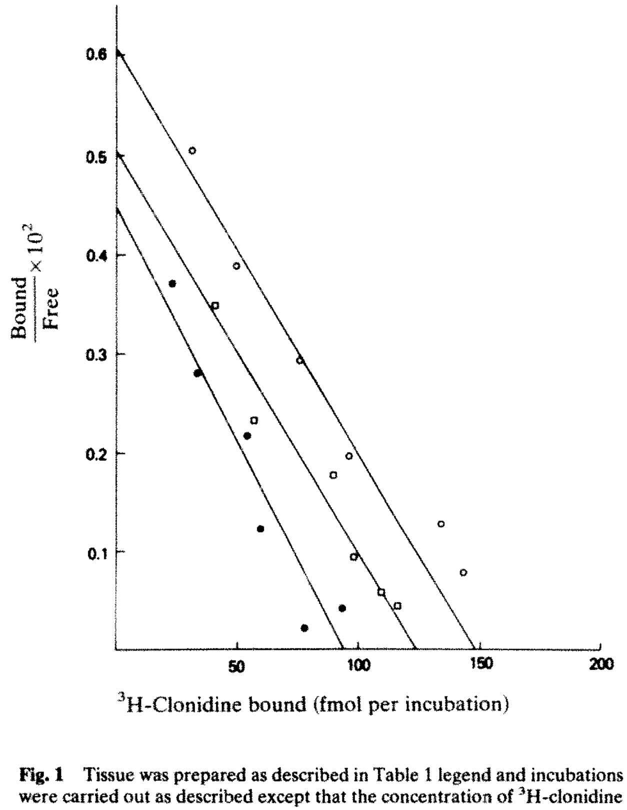


Fig. 1 Tissue was prepared as described in Table 1 legend and incubations were carried out as described except that the concentration of ^{3}H-clonidine was varied. Morphine treatment, either one or three pellet (over 2 days) implantation, resulted in significant increases in the apparent number of binding sites for clonidine. B_{max} values for control (●) 94 ± 8.0, for one pellet (○) 148 ± 7.7, and for three pellets (□) 127 ± 8.3. No attempt was made to measure the circulating level of morphine following these implants; however, randomly selected animals from each treated group showed the signs of withdrawal described in the text. The linear Scatchard plots could be converted to broken lines by including the modifiers mentioned in Table 1 in the incubation. In the presence of GTP (10^{-4}M), resolution of ligand binding data into two components with apparent affinities of 2.1 nM and 22.5 nM was possible. Similar results were obtained with tissue from chronic-morphine treated animals.

inetic reasons, shorter times of naloxone administration to morphine-treated animals were not examined. Chronic treatment with naloxone was not attempted.

Opiate-like substances are reported to act on brain noradrenergic neurones by (1) inhibiting the firing rates of the noradrenergic neurones of the locus coeruleus, (2) reducing noradrenaline levels and/or turnover, and (3) inhibiting the depolarization-induced release of noradrenaline from cerebral slices[7,14]. A prolonged suppression of noradrenergic activity caused by the inhibitory actions of morphine results in an increase in the apparent number of some noradrenergic target receptors, as reflected in the increased binding of ^{3}H-clonidine. Llorens *et al.*[13] have also reported that the binding of ^{3}H-DHA was significantly increased (by 19%) in rat cortical membranes. β-adrenergic binding data from our study confirmed this finding.

The demonstration that there were no changes in α-adrenergic binding using the α-antagonist WB-4101, whereas specific changes were apparent using ^{3}H-clonidine, presumably reflects occupancy of distinct populations of receptors by the two ligands. This supports the findings of other studies which report that ^{3}H-WB-4101 binds to one class of α-receptor and ^{3}H-clonidine to a second category demonstrating different biochemical and pharmacological characteristics[10,11].

The increase in binding of clonidine reported here contrasts with decreases in ^{3}H-clonidine binding sites following chronic treatment with the monoamine oxidase inhibitor clorgyline[15]. Additionally, we and others have found increases following intraventricular 6-hydroxydopamine[11]. It is interesting that both receptor changes are rapidly reversed in morphine-treated animals following abrupt withdrawal by naloxone injection, probably as a result of an increased turnover of noradrenaline[16]. In this regard, rapid desensitization of adrenergic receptors in brain slice experiments has been reported[17] and the time course of these changes is consistent with the changes reported here. However, decreases in clonidine binding following monoamine oxidase inhibitors shows a slower time course (days), which may reflect differences in noradrenaline release following the two treatments.

The distribution of opiate and α_2-adrenergic receptors shows some similarity in rat brain, overlapping particularly in areas implicated in opiate function[18,19]. It is possible, therefore, that some of the signs of withdrawal (such as clonidine-reversed increased noradrenaline release[16]) may be related to or amplified by the development of α_2-adrenergic supersensitivity during morphine treatment. The findings described here may be relevant to the effective use of clonidine in the treatment of withdrawal signs[20], and we are therefore exploring the possibility that similar receptor changes occur in humans following addiction.

Received 2 October 1980; accepted 25 February 1981.

1. Berthelsen, S. & Pettinger, W. A. *Life Sci.* **21** 595–606 (1977).
2. Gold, M.S., Redmond, D. E. Jr & Kleber, H. D. *Lancet* **ii**, 929–930 (1978).
3. Gold, M. S., Redmond, D. E. Jr & Kleber, H. D. *Am. J. Psychiat.* **136**, 100–101 (1979).
4. Cedarbaum, J. M. & Aghajanian, G. K. *Brain Res.* **112**, 413–419 (1976).
5. Redmond, D. E. Jr *et al. Brain Res.* **116**, 502–510 (1976).
6. Cedarbaum, J. M. & Aghajanian, G. K. *Eur. J. Pharmac.* **44**, 375–385 (1977).
7. Aghajanian, G. K. *Nature* **276**, 186 (1978).
8. Blasig, J., Herz, A., Rheinhold, K. & Zieglgänsberger, S. *Psychopharmacologia* **33**, 19–38 (1973).
9. Wei, E., Loh, H. H. & Way, E. L. *J. Pharmac. exp. Ther.* **184**, 389–403 (1973).
10. U'Prichard, D. C. & Snyder, S. H. *J. Neurochem.* **34**, 385–394 (1980).
11. U'Prichard, D. C., Bechtel, W. D., Rouot, B. M. & Snyder, S. H. *Molec. Pharmac.* **16**, 47–60 (1979).
12. Rosenblatt, J. E., Pert, C. B., Tallman, J. F., Pert, A. & Bunney, W. E. *Brain Res.* **160**, 186–101 (1979).
13. Llorens, C., Martres, M. P., Baudry, M. & Schwartz, J. C. *Nature* **274**, 603–605 (1978).
14. Redmond, D. E. Jr in *Animal Models in Psychiatry and Neurology* (eds Hanin, I. & Usdin, E.) (Pergamon, Oxford, 1977).
15. Campbell, I. C., Murphy, D. L., Gallager, D. W. & Tallman, J. F. in *Monoamine Oxidase: Structure Function and Altered Functions* (eds Singer, T. P., von Korff, R. W. & Murphy, D. L.) 517–530 (Academic, New York, 1979).
16. Laverty, R. & Roth, R. H. *Brain Res.* **182**, 482–485 (1979).
17. Maggi, A., U'Prichard, D. & Enna, S. J. *Science* **207**, 645–646 (1980).
18. Young, W. S. & Kuhar, M. J., *Proc natn. Acad. Sci U.S.A.* **77**, 1696–1700 (1980).
19. Young, W. S. & Kuhar, M. J. *Brain Res.* **179**, 255–270 (1979).
20. Uhde, T. W., Redmond, D. E. Jr & Kleber, H. D. *Psychiat. Res.* **2**, 37–47 (1980).

Contralateral denervation causes enhanced transmitter release from frog motor nerve terminals

Albert A. Herrera & Alan D. Grinnell

Department of Physiology, Jerry Lewis Neuromuscular Research Center, Ahmanson Laboratory of Neurobiology, School of Medicine, University of California, Los Angeles, California 90024, USA

It is becoming increasingly apparent that vertebrate neuromuscular junctions can exhibit considerable plasticity in transmitter release. Long-term changes in the level of release, and therefore the effectiveness of synaptic transmission, are seen in conditions of altered use[1,2], ageing[3,4], hibernation[5], disease[6], hormonal influence[7] and as a result of competitive interaction[8]. We have previously shown that transmitter release at frog neuromuscular junctions can vary over a wide range depending on the type of muscle examined[9,10] or motor unit size[11,12]. We now present evidence that unilateral denervation of the frog sartorius muscle causes a large increase in transmitter release from nerve terminals in the contralateral sartorius, without an increase in synaptic size. Consequently, the safety margin for neuromuscular transmission is enhanced[13]. Contralateral denervation also causes a significant increase in apparent poly-neuronal innervation; but in contrast to similar experiments with the frog cutaneous pectoris muscle[14–17], the increase in polyneuronal innervation does not seem to be due to nerve terminal sprouting.

The left sartorius muscle in 37 adult frogs (*Rana pipiens*, body length 4.5–7 cm, anaesthetized in 0.1% tricaine methanesulphonate) was completely denervated by one of the following procedures: (1) crushing of the sartorius nerve near its entrance into the muscle; (2) sartorius nerve crush with excision of the lateral or pelvic half of the sartorius muscle; (3) resection of a 1-cm length of spinal nerve 8 near the cord. The different operative techniques, designed for other studies involving transmitter release in reinnervated ipsilateral muscles[11], produced essentially the same changes in contralateral muscles. Synaptic effectiveness, the incidence of polyneuronal innervation and synaptic morphology were assessed 5–16 weeks later in the intact right sartorius as described previously[10,11] and below. Controls were normal sartorius muscles from 44 unoperated frogs, obtained from the same supplier at the same time. Muscles were immersed in Ringer solution containing 116 mM NaCl, 2 mM KCl, 1.8 mM $CaCl_2$, 1 mM $MgCl_2$ and about 1 mM $NaHCO_3$ (*p*H 7.2), or in Ringer containing reduced Ca^{2+} concentration as noted. Endplate potentials (e.p.ps) were

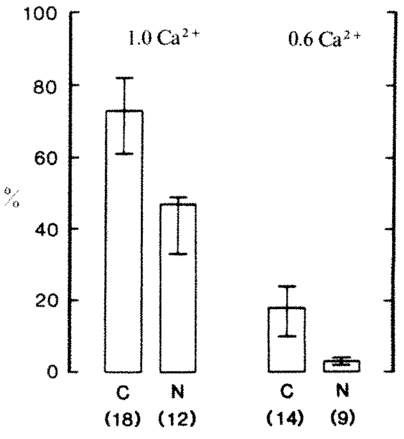


Fig. 1 The effect of lowered Ca^{2+} concentration (1.0 and 0.6 mM) on twitch tension evoked by nerve stimulation for sartorius muscles contralateral to a nerve crush (C) and normal sartorius muscles from unoperated frogs (N). Tension is expressed as a percentage of the tension generated in normal Ringer containing 1.8 mM Ca^{2+}. At each concentration the median, interquartile range and number of muscles tested (in parentheses) are shown.

0028-0836/81/240495—03$01.00

recorded intracellularly with standard techniques. The direct method of estimating quantal content, that is, mean e.p.p./mean miniature e.p.p., was used. Bath temperature (15 °C) and the degree of stretch of the muscle (1.1 × rest length) were carefully controlled. Data samples showing skewed distributions are described here by their medians and interquartile ranges (i.q.r.), and were compared with the Mann–Whitney U-test. Parametric statistics were used for other samples as indicated.

We first assessed synaptic efficacy by measuring the tension of twitches evoked by single nerve stimuli using Ringer containing various Ca^{2+} concentrations. When the Ca^{2+} concentration is lowered, tension falls as transmission at the strongest junction on some muscle fibres becomes subthreshold. Thus, at the particular low Ca^{2+} concentrations examined (1.0 and 0.6 mM), the extent to which tension has fallen relative to the tension generated in normal Ringer (1.8 mM) is a measure of safety margins at junctions in that muscle. Because many sartorius junctions have very low safety margins and others are even subthreshold for muscle contraction in normal Ringer[10,12], this procedure is a sensitive test for differences in synaptic effectiveness. Figure 1 shows that when Ca^{2+} concentration was lowered to 1 mM, twitch tension in contralateral muscles fell only to 73% (median of 18 muscles) of tension in normal Ringer. This was significantly more than tension remaining in normal muscles (47%, 12 muscles, $P < 0.002$). Similarly, in 0.6 mM Ca^{2+} residual twitch tension in contralateral muscles (18%, median of 14 muscles) was significantly higher than in normal muscles (3%, 9 muscles, $P < 0.002$). Similarly enhanced safety margins were seen at the longest post-operative time studied (111 days, 24% in 0.6 mM Ca^{2+}).

The second method of testing synaptic effectiveness was to measure transmitter release in terms of e.p.p. quantal content in Ringer containing 0.2 or 0.3 mM Ca^{2+} and 1 mM Mg^{2+}. In 0.2 mM Ca^{2+}, quantal content at junctions in contralateral muscles (median = 3.38, i.q.r. = 1.60–8.01, $n = 23$) was fivefold greater than at junctions in normal muscles (0.67, 0.21–1.75, $n = 21$), and in 0.3 mM Ca^{2+}, quantal content at contralateral junctions (median = 4.80, i.q.r. = 2.05–9.22, $n = 28$) was nearly threefold higher than at normal junctions (1.65, 0.97–2.87, $n = 45$). These differences were highly significant (0.2 mM Ca^{2+}: $P < 0.001$, 0.3 mM Ca^{2+}: $P < 0.001$).

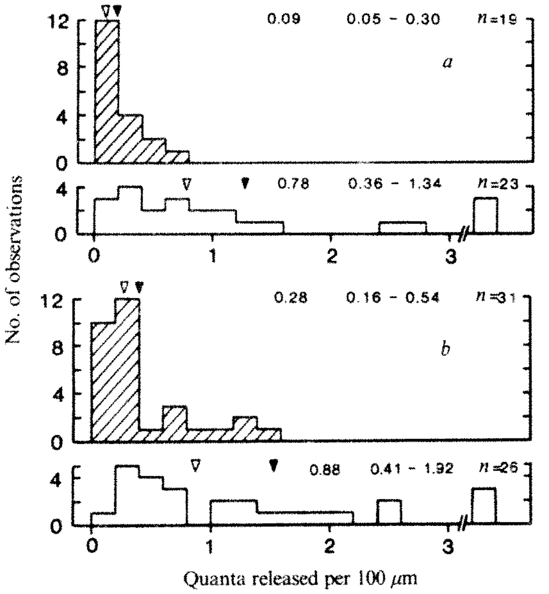


Fig. 2 Quanta released per 100 μm of nerve terminal length in Ringer containing 0.2 (a) or 0.3 mM Ca^{2+} (b) and 1 mM Mg^{2+}. In each case the upper (shaded) histogram shows data from junctions in normal unoperated sartorius muscles and the lower (unshaded) histogram data from muscles contralateral to a denervation. Figures shown are median (also indicated by ▽), interquartile range and sample size. Means are indicated by ▼. Actual data exceeding 3 quanta per 100 μm are 3.21, 4.49 and 5.13 in a and 4.84, 5.39 and 7.41 in b. Normalizing release to terminal length presupposes that release and terminal length are directly proportional. Such proportionality seems to hold for frog muscle when release is partially blocked by low Ca^{2+} concentration[19], but the correlation is marginal. In our experiments there was a significant positive correlation between quantal content and terminal length for normal sartorius junctions, but not for contralateral junctions. A non-correlation might be expected if release is highly non-uniform along the length of a terminal branch at low Ca^{2+} concentrations[25] (but see ref. 26). Such non-uniformity would affect consideration of possible physiological mechanisms of enhancement, but not the basic observation of greater release.

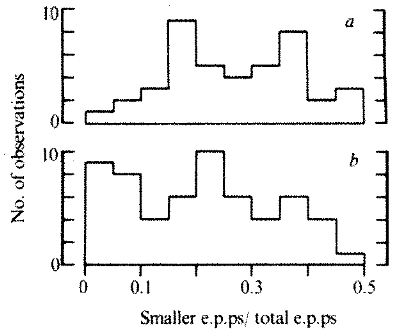


Fig. 3 For doubly innervated endplate sites, these histograms show the distributions of the ratios of the amplitude of the smaller e.p.p. component to total e.p.p. amplitude. a, Normal sartorius muscles, $n = 42$, mean ratio = 0.28; b, sartorius muscles contralateral to a denervation, $n = 58$, mean ratio = 0.22. Curare concentration $2\text{–}5 \times 10^{-6}$ M.

To determine whether this increase in transmitter release was due to an increase in synaptic size, contralateral and normal muscles were stained with the Karnovsky cholinesterase (ChE) stain[18]. As endplate size is proportional to muscle fibre diameter[19] we made sure that the muscle fibres bearing the endplates in the samples we compared had similar distributions of diameter (mean contralateral fibre diameter = 47 μm, mean normal fibre diameter = 46 μm). Endplate size, measured as total ChE-stained postsynaptic gutter length, was not significantly different for contralateral junctions (402 ± 11 μm s.e.m.; 333 endplates from 19 muscles) and normal junctions (416 ± 15 μm s.e.m.; 190 endplates from 8 muscles; $P > 0.40$). Thus, the enhanced quantal content at contralateral junctions apparently cannot be explained by an increase in synaptic size; but rather the inherent release properties of contralateral terminals seem to have been altered.

This conclusion is supported by measurements of the efficacy of the release mechanism in identified terminals. Junctions from which recordings were made were marked by intracellular dye injection, stained with nerve terminal (see below) and ChE stains, and measured in the light microscope. Quantal contents were then normalized to nerve terminal length and the result, expressed as quanta released per 100 μm of nerve terminal, can be viewed as a measure of the inherent transmitter release capability of a given terminal. Figure 2a shows that in 0.2 mM Ca^{2+}, median release per unit length was over eight times higher in contralateral muscles than in normal muscles (0.78 compared with 0.09), whereas in 0.3 mM Ca^{2+} (Fig. 2b) it was over three times higher (0.88 compared with 0.28). These differences were highly significant (0.2 mM Ca^{2+}: $P < 0.001$, 0.3 mM Ca^{2+}: $P < 0.005$). Mean values are also shown in Fig. 2 (closed arrows), but in such skewed distributions they describe the average less reliably. We found no evidence for any seasonal variation in any of our measurements of synaptic effectiveness.

In other intracellular measurements, we sought to estimate the incidence of focal polyneuronal innervation, defined as the convergence of two or more motor axons on to the same endplate site. Sartorius muscle fibres also show nonfocal polyneuronal innervation, that is, two or more well-separated endplate sites occur along the length of each fibre. Muscles were immersed in Ringer containing the minimum concentration of curare ($2\text{–}5 \times 10^{-6}$ M) necessary to block twitching. We then tried to subdivide the e.p.p. by changing the amplitude, polarity and duration of a stimulus to the nerve. This procedure, although customary, almost certainly underestimates polyneuronal innervation. In addition, we restricted our comparison of polyneuronal innervation to muscles tested at the same time of year, because there is evidence that the occurrence of terminal sprouting in normal frogs shows seasonal variation[20].

We found that the incidence of apparent polyneuronal innervation was higher in sartorius muscles contralateral to a denervation ($15\% = 59/392$ endplates from 11 muscles) than in normal sartorius muscles ($9\% = 21/227$ endplates from 6 muscles) when both types of muscles were tested at the same time of year (May–June). The difference between these proportions was significant ($P < 0.02$).

A search was made for morphological correlates of the higher incidence of focal polyneuronal innervation in contralateral muscles relative to normal muscles. Muscles were stained with the nitroblue tetrazolium (NBT) nerve terminal stain[21] and with ChE stain. NBT renders all nerve terminals and terminal sprouts clearly visible in the light microscope (Zeiss 40/0.75 water immersion objective, final magnification ×320). For example, the fine terminal sprouts which appeared commonly in the reinnervated sartorius muscles used in other studies[11] were consistently stained by the identical procedure. Careful examination of 30–50 well-exposed surface endplates in each of 10 contralateral muscles revealed no nerve terminal sprouts between synaptic sites on adjacent muscle fibres. In all other respects, contralateral junctions were indistinguishable from normal junctions in the light microscope. Thus, in contrast to reports for cutaneous pectoris muscles contralateral to a denervation[14,17], we find no morphological evidence in contralateral sartorius muscles either for nerve terminal elongation or terminal sprouting to innervate adjacent fibres. We cannot rule out the possibility that collateral (nodal) sprouts, or preterminal sprouts (branches from the unmyelinated portion of an axon proximal to the terminal) are responsible for the increased polyneuronal innervation. However, we see no terminal sprouts of the type described by Rotshenker and Reichert[14,17] in contralateral cutaneous pectoris muscles. In the absence of such evidence, an alternative explanation for the apparent increase in focal polyneuronal innervation can be suggested.

Our estimate of focal polyneuronal innervation in the normal sartorius is in the range of previous estimates of 6–30% (refs 22–24). Recently, however, we found that sartorius endplates may receive some inputs which are so weak that they may be easily obscured by the curare used to assess polyneuronal innervation[12]. If contralateral denervation causes enhanced release, some of the very weak polyneuronal inputs may become sufficiently strong to be detectable in curare. Figure 3 shows the distribution of the ratios of the amplitude of the smaller e.p.p. component relative to total e.p.p. amplitude at doubly innervated endplates sites in normal muscles (Fig. 3*a*) and contralateral muscles (Fig. 3*b*). The two histograms differ in that a greater proportion of the contralateral doubly innervated endplates receive one input which is much weaker than the other. Of the 58 doubly innervated contralateral endplates 17 (29%) show a ratio of e.p.p. components ≤ 0.1, whereas only doubly innervated normal endplates (7%) fall into this class. In the absence of any morphological evidence for terminal sprouting, it seems possible that the apparent increase in polyneuronal innervation at our contralateral endplates is due to the enhanced detectability of normally present polyneuronal inputs.

Further studies emphasized that for comparative purposes, it was important to control for seasonal variation in polyneuronal innervation. Normal sartorius muscles tested in mid-winter (January–February) showed enhanced polyneuronal innervation (18%, 56/311 endplates from 7 muscles). We are pursuing this seasonal difference.

Many questions remain unanswered regarding this contralateral effect. It is not clear, for example, whether the increase in transmitter release predominantly affects weak terminals as may be inferred from Fig. 3, or whether all terminals are involved. Further, our findings differ from those of Weakly and Yao[24], who found no change in polyneuronal innervation of frog sartorius junctions after contralateral denervation. We have no specific explanation for this discrepancy, except to suggest that polyneuronal innervation, as well as synaptic effectiveness and sprouting, are probably controlled in complex ways by several factors. We are now investigating the physiological mechanisms of the enhanced transmitter release, related effects in contralateral and ipsilateral synergists and antagonists, and the involvement of the spinal cord as a possible pathway for the effect. The results of these studies may have important implications for the interpretation of other experiments, using spinal isolation, deafferentation, tenotomy and local paralysis. It is clear, for example, that contralateral muscles may not serve as adequate controls for unilateral manipulations.

This work was supported by grants from the Muscular Dystrophy Association and from the NIH (NS06232) to A. D. G.; A. A. H. was supported by fellowships from the Muscular Dystrophy Association and the NIH.

Received 2 December 1980; accepted 27 March 1981.

1. Robbins, N. in *The Neurosciences, Third Study Program* (eds Schmitt, F. O. & Worden, F. G.) 953–960 (MIT Press, Cambridge, 1974); *Trends Neurosci.* **3,** 120–122 (1980).
2. Snider, W. D. & Harris, G. L. *Nature* **281,** 69–71 (1979).
3. Gutmann, E. & Hanzlíková, V. *Age Changes in the Neuromuscular System* (Scientechnica, Bristol, 1972).
4. Smith, D. O. *Expl Neurol.* **66,** 650–666 (1979).
5. Melichar, I., Brožek, G., Janský, L. & Vyskočil, F. *Pflügers Arch. ges. Physiol.* **345,** 93–106 (1973).
6. Cull-Candy, S. G., Miledi, R., Trautman, A. & Uchitel, O. D. *J. Physiol., Lond.* **299,** 621–638 (1980).
7. Tucek, S., Kostiova, D. & Gutmann, E. *J. neurol. Sci.* **27,** 353–362 (1976).
8. Mark, R. F. *Physiol. Rev.* **60,** 355–395 (1980).
9. Herrera, A. A. & Grinnell, A. D. *Brain Res.* **194,** 228–231 (1980).
10. Grinnell, A. D. & Herrera, A. A. *J. Physiol., Lond.* **307,** 301–317 (1980).
11. Herrera, A. A. & Grinnell, A. D. *Nature* **287,** 649–651 (1980).
12. Grinnell, A. D. & Herrera, A. A. *Soc. Neurosci. Abstr., 10th a. Meet.* Abstr. 192.11 (1980).
13. Herrera, A. A. & Grinnell, A. D. *Soc. Neurosci. Abstr., 10th a. Meet.* Abstr. 32.4 (1980).
14. Reichert, F. & Rotshenker, S. *Brain Res.* **170,** 187–189 (1979).
15. Rotshenker, S. *Israel J. med. Sci.* **14,** 1002 (1978); *Brain Res.* **155,** 354–356 (1978); *J. Physiol., Lond.* **292,** 535–547 (1979).
16. Rotshenker, S. & McMahan, U. J. *J. Neurocytol.* **5,** 719–730 (1976).
17. Rotshenker, S. & Reichert, F. *J. comp. Neurol.* **193,** 413-422 (1980).
18. Karnovsky, M. J. *J. Cell Biol.* **23,** 217–232 (1964).
19. Kuno, M., Turkanis, S. A. & Weakly, J. N. *J. Physiol., Lond.* **213,** 545-556 (1971).
20. Wernig, A., Pécot-Dechavassine, M. & Stöver, H. *J. Neurocytol.* **9,** 277–303 (1980).
21. Letinsky, M. S. & DeCino, P. A. *J. Neurocytol.* **9,** 305–320 (1980).
22. Vyskočil, F. & Magazanik, L. G. *Pflügers Arch. ges. Physiol.* **368,** 271–273 (1977).
23. Weakly, J. N. *Brain Res.* **158,** 235–239 (1978).
24. Weakly, J. N. & Yao, Y. M. *Brain Res.* **204,** 421–423 (1981).
25. Bennett, M. R. & Lavidis, N. A. *J. gen. Physiol.* **74,** 429–456 (1979).
26. De Cino, P. *J. Neurosci.* **1,** 308–317 (1981).

Ca-dependent K channels with large unitary conductance in chromaffin cell membranes

A. Marty

Max-Planck-Institut für biophysikalische Chemie, Am Faßberg,
D-3400 Göttingen, FRG

Injection of Ca ions increases the membrane permeability of many types of cells (reviewed by Meech[1]), and in molluscan neurones, removal of Ca ions from the external solution suppresses a voltage-dependent K conductance[2,3]. These two phenomena may be related through a potential-dependent Ca influx[2]. The question then arises whether the voltage dependence of the corresponding K permeability system is an intrinsic property of this system or whether it is merely a consequence of the increase of Ca influx with depolarization. The resolution of this and similar problems has been hampered by technical difficulties, due to the presence of voltage-dependent Ca channels, of Ca-independent K conductances and of strong intracellular buffering for Ca ions. In the present study patches of membranes containing functional, Ca-dependent K channels have been isolated from chromaffin cells. It is shown that application of low Ca concentrations to the inner side of the membrane affects the properties of the channels, while Ca ions are ineffective on the outer side of the membrane. The method used here overcomes several of the limitations of the previous studies using cellular recording.

0028–0836/81/240497—04$01.00

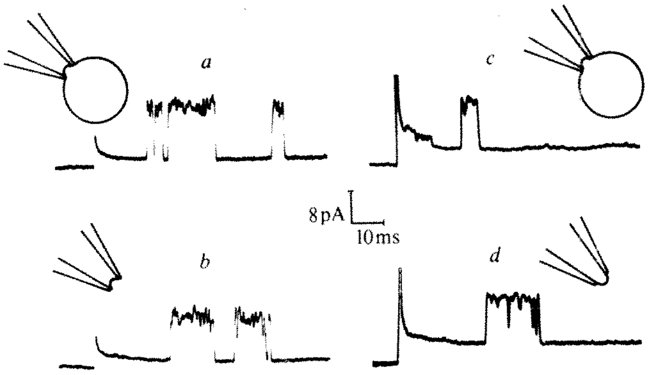


Fig. 1 Large K unitary currents in isolated patches. *a*, *b*, A pipette containing Ca-free Ringer (1 mM Co replacing Ca) was first used to record from a patch of a cell membrane (*a*). The bath contained a high-K solution (in mM: 143 KCl, 2 $MgCl_2$, 1 Ca-11 EGTA buffer, 10 HEPES-KOH, *p*H 7.3) which presumably brought the cell membrane potential close to 0 mV. The pipette was held at the bath potential, so that the holding potential for the patch was also 0 mV. *a*, 75-mV potential jumps were applied to the pipette interior to make the patch membrane potential positive. *b*, Response to the same voltage pulse after moving the pipette away from the cell (inside-out patch). Unitary outward currents are 12 pA in *a*, 10.5 pA in *b*. There is a pronounced noise increase during the opening of the channels. *c*, *d*, A pipette containing a high-K solution (same as the bath solution in *a*, *b*) was first applied to a cell bathed in normal Ringer (*c*). The pipette interior was held at the bath potential so that the patch holding potential was equal to the cell resting potential (presumably ~−60 mV). 100-mV positive pulses activate large (11 pA) K unitary currents. *d*, Similar currents are obtained after destruction of the initial patch and withdrawal of the pipette when applying 104-mV pulses from a holding potential of −64 mV (outside-out patch). 1 kHz filtering.

Isolated bovine chromaffin cells[4] were cultured for up to 8 days in medium 199 (Gibco) supplemented with 10% fetal calf serum, bovine serum albumin (1 mg ml^{-1}) and antibiotics. They remained in good condition as judged by their ability to grow dendrite-like structures, to maintain a high resting potential and to fire action potentials. High-resolution measurements of currents were obtained through a glass micropipette tightly sealed onto a patch of membrane ('gigaseal' method[5-8]). The potential in the pipette was under experimental control, while the potential on the other side of the membrane was either the cell resting potential or, in isolated patches, the bath potential. In all experiments, a large-resistance (>10 GΩ) seal was established between the pipette and the membrane of the intact cell. Large unitary outward currents were then obtained on depolarizing the patch membrane (Fig. 1*a*, *c*). The corresponding channels will be fully characterized elsewhere; they are K selective (as indicated by a reversal potential value more negative than −60 mV in normal Ringer and of about 0 mV in isotonic K solution) and have a very high unit conductance (~180 pS in symmetrical K). The frequency of the current pulses at a given potential is strongly reduced if Ca ions are omitted from the pipette solution. Compared with the Ca-dependent single K-channel recordings recently described in *Helix* neurones[9], they differ by having a higher unit conductance and by the lack of a delay in turn-on on large depolarizing voltage commands.

After forming a seal, cell-free patches of membrane were isolated in two ways. In one type of experiment, the pipette was simply moved away from the cell in a low-Ca, high-K medium to yield an 'inside-out patch'[6-8] (Fig. 1*a*, *b*). Positive potential jumps from 0 mV activated large outward unitary currents with very similar properties before and after withdrawal of the pipette from the cell. (Voltage differences are given with respect to the patch membrane: the potential on the outer side is taken as zero. Current directions are also referred to the patch, with outward membrane currents being displayed upwards in the records.) In the other type of experiment, the initial patch was destroyed by applying suction to the pipette interior and the pipette was then withdrawn. In most cases it then sealed over with a new patch of membrane pulled away from the cell[8] ('outside-out patch', Fig. 1*d*). Whereas inside-out patches have the internal (previously intracellular) side of the membrane in contact with the bath, outside-out patches have their external side exposed to the bath[8].

Initial attempts to change the bath solution with various Ca-EGTA buffers in the conditions of Fig. 1*b*, *d* led to erratic results, probably because vesicular structures remained at the pipette tip instead of the desired single membrane patches[5]. These vesicular structures were avoided in later experiments. In the inside-out configuration (Fig. 1*b*), the outer part of the vesicle was destroyed chemically (Ca-EGTA buffer) and, if necessary, mechanically (by exposing briefly the pipette tip to air, or by bringing it in contact with a hexadecane droplet formed at the tip of a capillary glass tube held in the bath). None of these methods has a 100% success rate; the experiment was not continued until there was good evidence that a genuine inside-out patch had been obtained. (In cell-free conditions, the channels should have a unit size and kinetic properties virtually identical to those observed before leaving the cell, as illustrated in Fig. 1*a*, *b*.) In the outside-out configuration (Fig. 1*d*), the time constant of the whole-cell membrane capacitive current was measured when the initial seal had been destroyed by suction and the pipette was still in contact with the cell. Low values (<0.2 ms) of this time constant ensured a low (<20 MΩ) resistance through the initial patch which could then be considered as effectively destroyed. With these precautions, the exchange of the bath solution (in the conditions of Fig. 1*b*, *d*) gave fast, stable and reliable results on the channel properties. The bath contained about 1 ml of solution, and was continuously perfused at a rate of ~1 ml min^{-1}. Normal Ringer contained (in mM): 140 NaCl, 2.8 KCl, 2 $MgCl_2$, 1 $CaCl_2$, 10 HEPES-NaOH, *p*H 7.2–7.3. The solutions used on the inner side of the membrane contained 143 KCl, 2 $MgCl_2$, 10 HEPES-KOH, *p*H 7.2–7.3, and a Ca-EGTA buffer. The effective free Ca concentration, Ca_i, was calculated using an apparent dissociation constant of 10^{-7} M (estimated from ref. 10). All experiments were carried out at room temperature.

If the Ca concentration is altered at the external side of an outside-out isolated patch (Fig. 1*d* situation), no conspicuous effect is seen on the K channels. An example of such an experiment is illustrated in Fig. 2.

In sharp contrast, Fig. 3 shows the striking effects that follow a change in Ca-EGTA buffers on the internal side of an inside-out

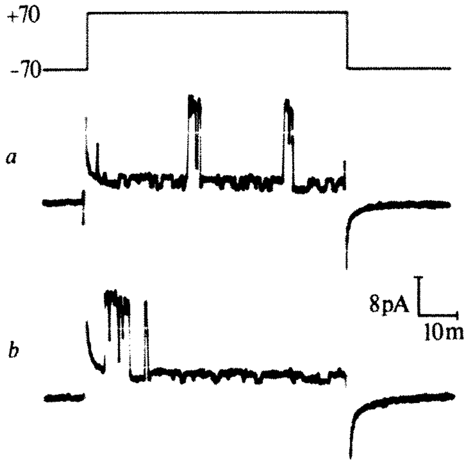


Fig. 2 Lack of effect of external Ca in an outside-out patch. 140-mV positive pulses were applied from −70 mV in an outside-out patch bathed in normal Ringer (*a*) and in 0 Ca, 1 mM Co Ringer (*b*). The pipette contained a high-K solution and a Ca-EGTA buffer with a calculated Ca concentration of 2×10^{-7} M. The voltage jumps elicit large (15 pA) K currents as well as smaller outward currents (see also Fig. 1*c*) not studied here. Suppressing the external Ca ions does not affect the large K currents. 2 kHz filtering.

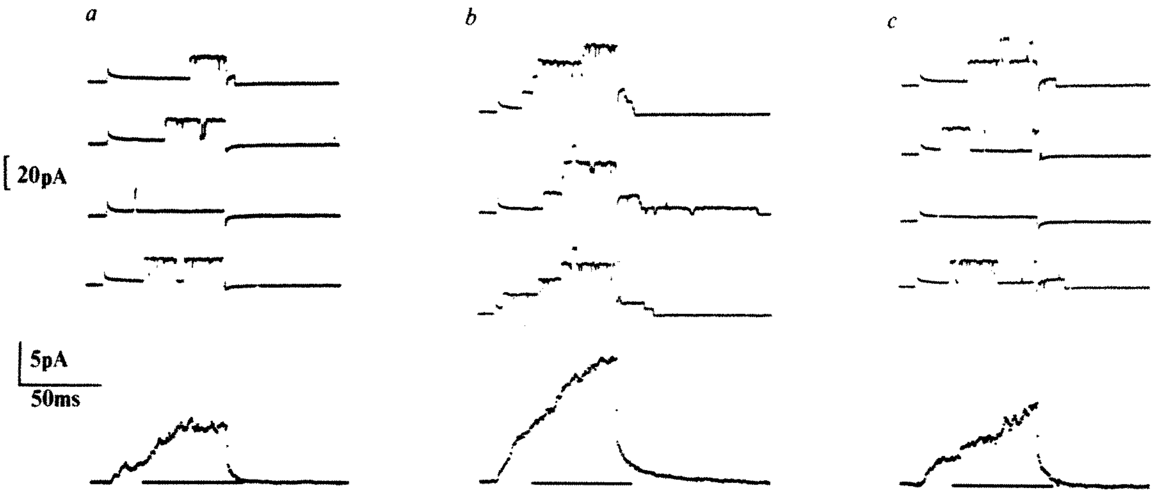


Fig. 3 Effects of various Ca-EGTA buffers in an inside-out patch. An inside-out patch was exposed to a high-K solution containing 1 mM Ca+ 11 mM EGTA (*a*, *c*) and 1 mM Ca+1.5 mM EGTA (*b*). Calculated free Ca concentrations are 10^{-8} M in *a* and *c*, and 2×10^{-7} M in *b*. 67-mV positive potential jumps from a holding potential of 0 mV were applied at ~1 Hz. The pipette contained Ca-free Ringer (1 mM Co replacing Ca). *a*, Four records in the initial low-Ca solution (upper traces). The unitary current is 4.5 pA at 0 mV, and 13.4 pA at +67 mV. Lower trace: average of 50 consecutive sweeps (the background current has been subtracted). *b*, Higher Ca solution. The unitary current is 4.0 pA at 0 mV and 9.6 pA at +67 mV. The number of open channels is larger during the pulse than in *a* and there is a marked lengthening of the tail of activity after returning to 0 mV. These two features are also evident in the average of 47 consecutive sweeps (lower trace). *c*, Records taken ~1 min after returning to the initial solution. The unitary current is 5.0 pA at 0 mV and 13.2 pA at +67 mV. The kinetic properties of the currents are almost as in *a* (lower trace, 37 consecutive sweeps). 1 kHz filtering. Vertical scale: 20 pA for single sweeps (upper traces) and 5 pA for average currents (lower traces). Baseline is indicated by horizontal bars in average current traces.

patch (Fig. 1*b* situation). The buffers used had estimated free Ca concentrations of 10^{-8} M and 2×10^{-7} M. Positive potential steps of 67 mV were applied from 0 mV. The elementary current is lower at the higher Ca concentration, with ratios of 0.89 at 0 mV and 0.72 at 67 mV. The average number of open channels is 0.51 at the end of the positive pulse and $\sim5\times10^{-4}$ before the pulse in 10^{-8} M Ca. This contrasts with 1.7 and 0.02 respectively in 2×10^{-7} M Ca. These values show that Ca increases the average number of open channels. All effects were readily reversible (Fig. 3*c*).

Figure 3 typifies results obtained with Ca-EGTA buffers having calculated Ca_i values of 10^{-9} M, 10^{-8} M (Fig. 3*a*), 5×10^{-8} M, and 2×10^{-7} M (Fig. 3*b*). All Ca effects illustrated in Fig. 3 were gradually observed on this concentration scale. In control experiments, the effects of two Ca-EGTA buffers having the same calculated Ca_i value but different EGTA concentrations were compared. These buffers (in mM: 11 EGTA-1 Ca and 1.1 EGTA-0.1 Ca) had identical effects on the K channels indicating that EGTA *per se* is ineffective. This result contrasts with a previous report on perfused snail neurones[11].

Figure 4 illustrates the result of changing the bath solution to one containing 1 mM Ca (and no EGTA). The unitary conductance is decreased, the mean number of open channels greatly increased and the voltage sensitivity entirely lost. The channels have open times of ⩾1 s (not illustrated). Again, all effects are reversible. Furthermore, application of 1 mM Ca_i to inside-out patches at membrane potentials of ⩾+60 mV resulted in multiple interruptions (of ~1 ms each) of the single channel currents, like those observed with QX222 on acetylcholine-sensitive channels[12]. These interruptions are probably due to a

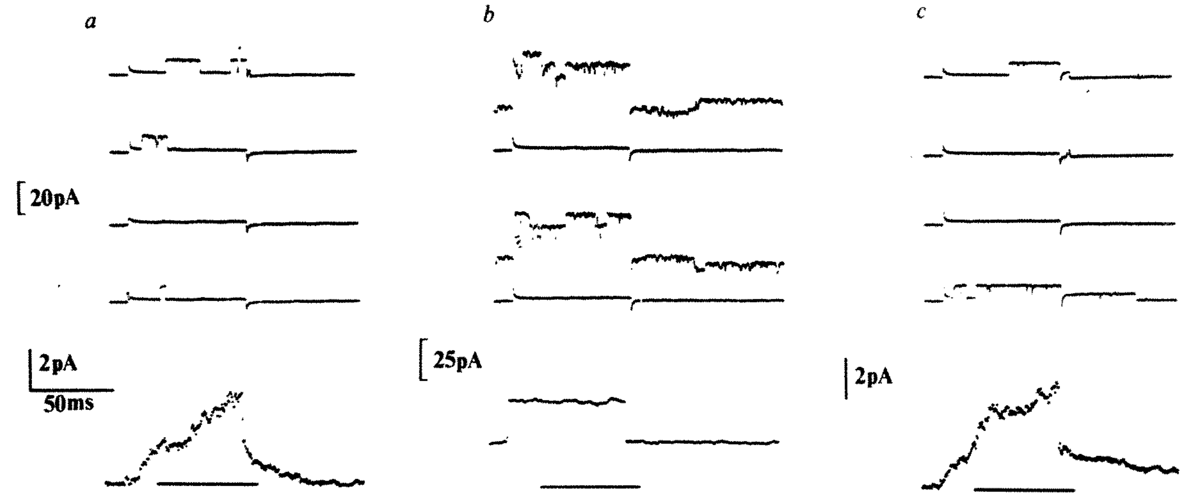


Fig. 4 Effects of 1 mM Ca in an inside-out patch. 40-mV positive potential jumps were applied from a holding potential of 0 mV to the same patch as in Fig. 3. *a*, Same buffer as in Fig. 3*a*. *b*, The bath was changed to a 1 mM Ca–0 mM EGTA solution. *c*, After returning to the initial buffer (recordings start 30 s after the solution change). Single sweeps (upper traces) and averaged currents (lower traces; 51, 21 and 48 consecutive sweeps in *a*, *b* and *c*, respectively) are illustrated. The background currents in the two upper traces in *b* were taken shortly after returning to the EGTA buffer. The unitary currents are at 0 mV, 4.2 pA (*a*), 3.8 pA (*b*) and 4.0 pA (*c*); at +40 mV, they are 8.0 pA, 7.2 pA and 7.6 pA, respectively. Statistical analysis of the traces in *b* suggests that 10 active channels were present in the patch. This value makes the probability of channel opening to be 0.76 in the presence of 1 mM Ca. (The probability of opening in the other traces of Figs 3 and 4 can be similarly calculated as the mean number of open channels divided by 10.) 1 kHz filtering. Vertical scale: 20 pA for single sweeps; 2 pA(*a*, *c*) and 25 pA(*b*) for average currents. Baseline indicated by horizontal bars in average current traces.

blockade of the channels by internal Ca ions. Previous observations of a reduction of K currents after injection of Ca ions into molluscan neurones[3] may be, at least in part, related to such an effect.

The present work is the first report of Ca-dependent K channels in chromaffin cells. These channels display an unusually large unitary conductance, and their properties are affected by the Ca concentration on the inner side of the membrane. In intact cells this concentration may be locally altered by voltage-dependent inflow of Ca ions as suggested by Meech and Standen[2]. Clearly, the action of Ca does not require any other soluble cytoplasmic substance.

For low internal calcium concentrations, the Ca-dependent K channels of chromaffin cells are voltage dependent. However, their sensitivity to the membrane potential is affected by Ca, and at 1 mM inner Ca concentration no voltage dependence is observed. Some investigators[13,14] have found that the K conductance elicited by injections of Ca ions to molluscan neurones is virtually voltage independent, whereas others have found a voltage-dependent conductance[15,16]. The present work suggests that the former results were simply obtained with larger concentrations than the latter.

Finally, the results show that the unitary conductance of K channels depends on Ca$_i$. This variation is already observable for very low Ca concentrations. Probably some common mechanism underlies this phenomenon as well as the effects of Ca$_i$ on the channel opening probability and its voltage dependence.

I thank Drs E. Fenwick and E. Neher for preparing and culturing the cells, Drs P. Ascher and F. Sigworth for comments on the manuscript, and Dr Neher for valuable advice at all stages of the work. This research was supported by the Deutsche Forschungsgemeinschaft and the Centre National de la Recherche Scientifique.

Received 29 January; accepted 9 April 1981.

1. Meech, R. W. *A. Rev. Biophys. Bioengng* **7,** 1–18 (1978).
2. Meech, R. W. & Standen, N. B. *J. Physiol., Lond.* **249,** 211–239 (1975).
3. Heyer, C. B. & Lux, H. D. *J. Physiol., Lond.* **262,** 349–382 (1976).
4. Fenwick, E. M., Fajdiga, P. B., Howe, N. B. S. & Livett, B. G. *J. Cell Biol.* **76,** 12–30 (1978).
5. Neher, E. in *Techniques in Cellular Physiology* (ed. Baker, P. F) (Elsevier, Amsterdam, 1981).
6. Horn, R. & Patlak, J. Proc. *natn. Acad. Sci. U.S.A.* **77,** 6930–6934 (1980).
7. Hamill, O. P. & Sakmann, B. *J. Physiol., Lond.* **312,** 41P–42P (1981).
8. Hamill, O. P., Marty, A, Neher, E., Sakmann, B. & Sigworth, F. J. *Pflügers Arch. ges. Physiol.* (submitted).
9. Lux, H. D., Neher, E. & Marty, A. *Pflügers Arch. ges. Physiol.* **289,** 293–295 (1981).
10. Caldwell, P. C. in *Calcium and Cellular Function* (ed. Cuthbert, A. W.) (Macmillan, London, 1970).
11. Kostyuk, P. G. & Krishtal, O. A. *J. Physiol., Lond.* **270,** 569–580 (1977).
12. Neher, E. & Steinbach, J. H. *J. Physiol., Lond.* **277,** 153–176 (1978).
13. Meech, R. W. *J. Physiol., Lond.* **237,** 259–277 (1976).
14. Hofmeier, G. & Lux, H. D. *Proc.* 28*th int. Congr. physiol. Sci.*, Budapest (Akademiai Kiado, in the press).
15. Ascher, P., Marty, A. & Neild, T. O. *J. Physiol., Lond.* **278,** 177–206 (1978).
16. Gorman, A. L. F. & Thomas, M. V. *J. Physiol., Lond.* **308,** 287–313 (1980).

Monoclonal antibodies against human fibroblast interferon

H. K. Hochkeppel, U. Menge & J. Collins

Gesellachaft für Biotechnologische Forshung mbH, Mascheroder Weg 1, D-3300 Braunschweig, FRG

The advantages of using monoclonal antibodies account for current interest in antibodies specific for interferon[1-3]. The successful isolation of a hybridoma line secreting mouse monoclonal antibodies to human leukocyte interferon (IFN-α) has already been reported[4]. We have now established a stable hybrid myeloma line (FO) which secretes antibody against human fibroblast interferon (IFN-β). We have demonstrated that, even after removal of much of the sugar moiety of IFN-β, the monoclonal anti-IFN-β antibodies neutralize the antiviral activity of human IFN-β but not that of human IFN-α. We suggest the possible use of these monoclonal antibodies for the identification and purification of IFN-β expressed by DNA cloned in bacteria.

Table 1 Inhibitory effect of anti-IFN-β antibody preparations on IFN-β antiviral activity

Anti-IFN-β antibody preparation	50% inhibition of IFN-β antiviral activity
Culture supernatant	1:7 dilution
Ascities	1:70 dilution

Various dilutions of monoclonal antibody were tested against IFN-β in the direct antiviral interferon assay on human FS4 fibroblasts with vesicular stomatitis virus (1,000 plaque forming units per ml) as challenge. The antibody titrations were carried out using 40 μl of diluted antibody per well of the microtitre plate. Three units of IFN-β were used for each antibody sample. Ascites tumours were obtained from mice pretreated with Freund's adjuvants injected with 10^6 hybridoma cells into the peritoneal cavity. After 2–3 weeks ascites fluids were collected and cells removed by centrifugation. Crude IFN-β was purified using an aqueous two phase polyethyleneglycol phosphate system[10]. Human IFN-β enters the upper phase while the bulk of the protein is found in the lower phase. The yield was 74–94%, corresponding to $3–7.3 \times 10^6$ units per mg protein.

One μg of purified blue Sepharose adsorbed IFN-β (ref. 5) was injected each week into 8-week-old female BALB/c mice; those having an anti-IFN antibody titre that inhibited human IFN-β in a direct antiviral assay[6] at a serum dilution of 1:200 were chosen for lymphocyte fusion experiments. A final booster injection of 1 μg IFN-β with the sugar moiety removed by treatment with a mixture of sugar-degrading enzymes was administered 4 days before spleen excision.

Myeloma FO[7] and spleen cells were fused using the method of Köhler and Milstein[1] with the following modifications. First, a feeder layer was omitted; instead the hybrids were grown in HAT (hypoxanthine–aminopterin–thymidine) medium containing 1% methylcellulose. Individually growing hybridoma colonies were picked up with sterile capillaries after 2 weeks and cloned separately in Costar tissue plate wells. The hybridoma colonies were then tested for inhibition of antiviral activity of IFN-β on FS4 fibroblasts in an indirect antiviral assay[4]. Positive cultures were subcloned, grown for 4 days and tested using an enzyme linked immunosorbent assay according to Towbin *et al.*[8] and a direct antiviral IFN-β assay. Cultures which were positive for antibody secretion in both the enzyme linked immunosorbent assay (ELISA) and the antiviral assay were subcloned once more and the supernatant of these again tested in the direct antiviral IFN-β assay[6]. In this way a stable hybridoma line secreting monoclonal IFN-β antibodies was established and checked for constant antibody secretion over a period of several months.

Inhibition of the antiviral activity of human IFN-β with monoclonal antibodies from hybridoma Costar tissue plate cultures was tested using increasing dilutions of the antibody sample. Partial inhibition of the antiviral activity of IFN-β was observed with antibody diluted by as much as 1:20. After treatment of the antibody sample with trypsin (10 μg ml^{-1}) for 60 min at 37 °C and subsequent inhibition of trypsin with soybean trypsin inhibitor, the inhibitory effect of the antibody sample on IFN-β in the direct antiviral assay was completely lost, which proves that the inhibition of the antiviral interferon activity was due to a protein. In addition, to determine whether the monoclonal antibodies were directed against the sugar or the protein moiety, human IFN-β was treated with a mixture of sugar-degrading enzymes (neuramidase, α-glucosidase, α-mannosidase, α/β-galactosidases) and this interferon sample used as standard IFN-β in the direct antiviral assay. The antibodies seem to be directed against the protein part of IFN-β as the antiviral activity of the treated sample was still completely inhibited by the antibodies at the same dilution. It was shown that monoclonal anti-IFN-β antibodies do not cross-react with human IFN-α (from 'buffy coat'). Aliquots of immunoglobulin-secreting hybridoma cells were injected into the peritoneal cavity of female BALB/c mice which had been pretreated with Freund's adjuvants 1 week before to grow ascites. After 1 week, ascites fluid was collected from the peritoneal cavity containing

0028–0836/81/240500—02$01.00

the monoclonal antibodies. This sample was found to have a 10-fold higher titre against standard human IFN-β in a direct antiviral interferon assay when compared with supernatant of the original monoclonal cultures (Table 1).

In an additional assay partially purified IFN-β was subjected to SDS-gel electrophoresis using the method of Laemmli[9] in parallel with molecular weight protein standards and then electrophoretically blotted on to a nitrocellulose sheet by a method described elsewhere[8]. While the standard proteins were stained with Coomassie brilliant blue the blotted interferon sample was assayed in an ELISA test developed for protein blots, using monoclonal anti-IFN-β antibodies (Fig. 1). Figure 1 shows that the antibodies specifically react with the protein band possessing a molecular weight (MW) of ~ 19,000. Monoclonal anti-IFN-β antibodies which were initially purified by 35% ammonium sulphate precipitation followed by passage over a Sephadex G-200 column were coupled to 2 ml of cyanogen bromide-activated Sepharose (Pharmacia) and tested for specific immunoadsorption of IFN-β. Figure 2 demonstrates that the antibodies were indeed specific for IFN-β.

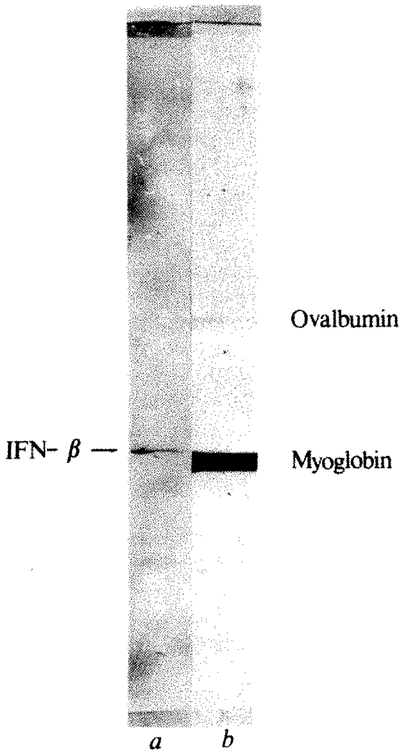


Fig. 1 Electrophoretic transfer of IFN-β from an SDS gel onto nitrocellulose sheet and subsequent ELISA reaction. *a*, IFN-β (ELISA); *b*, molecular weight standards stained with amino black. Electrophoresis was performed according to Laemmli[9]. The gel was then soaked in buffer containing 50 mM NaCl, 2mM Na-EDTA, 4 M urea, 0.1 mM dithiothreitol, 10 mM Tris-HCl, *p*H 7.0, according to Bowen *et al.*[12] and the proteins (IFN-β and markers) were electrophoretically transferred onto a nitrocellulose sheet[8]. While the section of the blot containing the markers was stained with amido black, the part containing IFN-β was soaked for 4 h at 37 °C in 3% bovine serum albumin (BSA)/Tris-buffered saline to saturate all nonspecific protein-binding sites for ELISA. After washing, the monoclonal antibody was added in a 1:50 dilution in 5% horse serum/0.5% BSA for 16 h at room temperature. The blot was then soaked for 2 h in a rabbit antimouse peroxidase-conjugated immunoglobulin G (IgG) solution (1:500 dilution in 5% inactivated horse serum/0.5% BSA), washed again then soaked for a further 2 h in a goat anti-rabbit peroxidase-conjugated IgG solution (1:500 dilution as above) to make the ELISA reaction more sensitive. After a final wash, the colour reaction was performed with *O*-dianisidin as substrate[6]. Marker proteins were ovalbumin, 45,000 (45K) and myoglobin, MW 17.5 K.

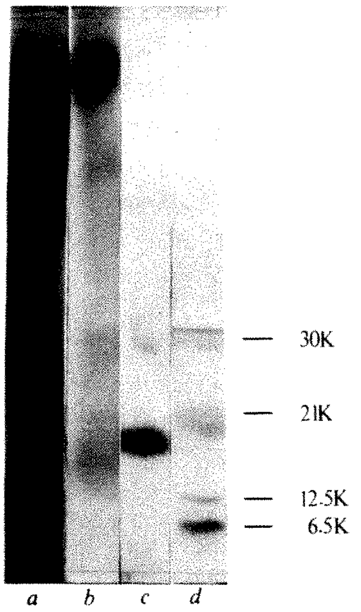


Fig. 2 Autoradiograph of an SDS gel with *a*, unpurified ^{125}I-labelled IFN-β (4-day exposure); *b*, unpurified ^{125}I-labelled IFN-β (1-day exposure); *c*, purified ^{125}I-labelled IFN-β (4-day exposure); *d*, protein markers. Monoclonal antibodies against IFN-β were purified from ascities fluid by a 35% ammonium sulphate precipitation followed by separation over a Sephadex G-200 column in a Tris–glycine buffer, *p*H 7.6, containing 0.04% *n*-butanol. The purified antibodies were then coupled to cyanogen bromide-activated Sepharose 4-B. An IFN-β sample, iodinated with ^{125}I according to Hunter and Greenwood [3], was purified over the column. A neutral wash with Tris-buffered saline at *p*H 7.4 removed nonspecifically adsorbed protein material. Specifically bound ^{125}I-labelled IFN-β was then washed off the column with 0.1 M glycine–HCl buffer, *p*H 2.3. The eluted sample was lyophilized, and then electrophoresed on an SDS gel according to Laemmli[9] parallel with unpurified ^{125}I-labelled IFN-β and ^{14}C-labelled protein markers. After electrophoresis the gel was soaked twice for 10 min in dimethyl sulphoxide (DMSO), and then in DMSO/30% diphenyloxazole for 20 min, washed with H_2O and dried. The dried gel was autoradiographed.

In an ELISA microtitre assay (not shown), anti-IFN-β monoclonal antibodies did not cross-react with a total protein preparation of *Escherichia coli.* Thus, as the monoclonal antibodies are probably directed specifically against the protein part of IFN-β, they could perhaps be used for the identification and purification of IFN-β expressed by DNA cloned in bacteria.

It is possible that our two stable hybridoma lines which secrete monoclonal anti-IFN-β antibodies are directed against different determinants of the IFN-β protein moiety. Note that there were fluctuations in the inhibitory activity of the monoclonal antibodies, between antiviral interferon assays which may be either due to the relative insensitivity of the assay system itself or the existence of more than one human IFN-β gene[10,11].

We thank M. Fitzek and K. Littmann for technical assistance, the Rentschler company for purified human IFN-β isolated from poly [r(I)]·poly[r(C)]-induced FS4 cells and H. Mohr for purified human IFN-α isolated from buffy coats.

Received 17 January; accepted 24 March 1981.

1. Köhler, G. & Milstein, C. *Nature* **256,** 495–497 (1975).
2. Köhler, G. & Milstein, C. *Eur. J. Immun.* **6,** 511–519 (1976).
3. Köhler, G., Hengartner, H. & Shulman, M. J. *Eur. J. Immun.* **8,** 82–88 (1978).
4. Secher, D. S. & Burke, D. C. *Nature* **285,** 446–450 (1980).
5. Menge, U. & Kula, M. R. German patent application No. P 2943016.8.
6. Havell, E. A. & Vilcek, J. *Antimicrob. Agents Chemother.* **2,** 476–484 (1972).
7. de Fazekas, S. E., Groth, & Scheidegger, S. *J. immun. meth.* **35,** 1–25 (1980).
8. Towbin, H., Staehelin, T. & Gordon, J. *Proc. natn. Acad. Sci. U.S.A.* **76,** 4350–4354 (1979).
9. Laemmli, U. K. *Nature* **227,** 680–685 (1970).
10. Weissenbach, J. *et al. Proc. natn. Acad. Sci. U.S.A.* **77,** 7152–7156 (1980).
11. Sagar, A. D., Pickering, L. A., Sussman-Berger, P., Stewart W. E. II & Sehgal, P. B. *Nucleic Acids Res.* **9,** 149–160 (1981).
12. Bowen, B., Steinberg, J., Laemmli, U. K. & Weintraub, H. *Nucleic Acids Res.* **8,** 1–19 (1980).
13. Hunter, W. M. & Greenwood, I. C. *Nature* **194,** 495–496 (1962).

Production of both macrophage activating and inhibiting activities by a murine T-lymphocyte hybridoma

C. Michael Jones, James A. Braatz
& Ronald B. Herberman

Laboratory of Immunodiagnosis, National Cancer Institute,
Bethesda, Maryland 20014, USA

T lymphocytes are known to secrete a variety of factors which affect macrophage function[1], including those which inhibit migration (MIF) and activate macrophages (MAF) to become cytotoxic[1], and produce neutral proteases[2], complement components[3] and other factors. Physical separation of MIF from MAF has not been achieved[4] and efforts directed towards characterizing these factors have been limited chiefly by their generation from lymph nodes or spleen cells and the difficulty in performing most macrophage function assays. We elected to approach the problem of generation of lymphokine by fusing concanavalin A (Con A)-activated AKR spleen cells to the AKR T-cell lymphoma BW5147 and measuring the biological activity using established assays. We describe here the generation of a cloned T-cell hybridoma whose supernatants produce both inhibition of macrophage migration and activation of macrophages for the production of C2 and elastase.

The hypoxanthine/aminopterin/thymidine (HAT) sensitive variant of BW5147 was fused using polyethylene glycol (Sigma, molecular weight 1,000) to Con A-activated AKR spleen cells[5]. Resultant hybridomas were divided into 51 2-ml aliquots and grown in the presence of HAT-selective medium. Thirty days post-fusion only five colonies exhibited macroscopic growth. The supernatants from these colonies were screened first for inhibition of macrophage migration (Table 1); control supernatants were collected from BW5147 growing in the same medium but lacking aminopterin at 5×10^5 cells ml^{-1}. The supernatants from two colonies significantly inhibited macrophage migration in the agarose droplet microassay but three did not. A positive control obtained with supernatant from Con A-stimulated spleen cells is shown for comparison (Table 1). These same colonies were screened for their ability to enhance the production of the second complement component by macrophage monolayers[3]. W25 and W38 enhanced the production of C2 4- and 2.5-fold, respectively, as shown in Table 1. Unfortunately, W25 was unstable and died. W38 was cloned by limiting dilution without feeder layers, resulting in the generation of four clones which were subjected to the same screening procedures.

Table 1 Inhibition of macrophage migration and enhancement of macrophage C2 production induced by T-hybridoma supernatants

	Colonies			Clones of W38	
	Migration index	Units per ml	No.	Migration index	Units ml^{-1}
		C2			C2
W38	0.55	0.24	1	0.62	0.78
W25	0.68	0.38	2	0.97	0.31
W31	1.08	0.12	3	0.85	0.35
W8	1.36	0.11	4	0.87	0.48
W32	1.29	0.145			
BW5147		0.095			
AKR spleen cell medium					0.12
AKR spleen cell pulsed with 6 μg ml^{-1} Con A				0.74	0.28

Inhibition of migration of C57BL/6 oil-induced macrophages, as measured by the agarose droplet technique[7]. Migration was performed in a 1:1 (colonies) or 1:4 (clones) dilution of test medium in RPMI 1640 and migration areas were measured at 20 h with a Zeiss inverted microscope. Results are expressed as the migration index, using the migration area given with BW5147 supernatants as the denominator. Final test volume per well was 0.2 ml. Production of C2 by oil-induced C57BL/6 macrophage monolayers, stimulated with supernatants from T-cell hybridomas. Macrophages at a concentration of 10^7 ml^{-1} in RPMI 1640 containing 10% fetal calf serum were allowed to adhere overnight to Costar 2-ml 'cluster wells' and then washed free of non-adherent cells. A 1:10 dilution of test medium in RPMI 1640 was added and aliquots were removed after 60 h of incubation and assayed for C2 by the haemolytic technique of Borsos *et al.*[8]. Final test volume per well was 2 ml.

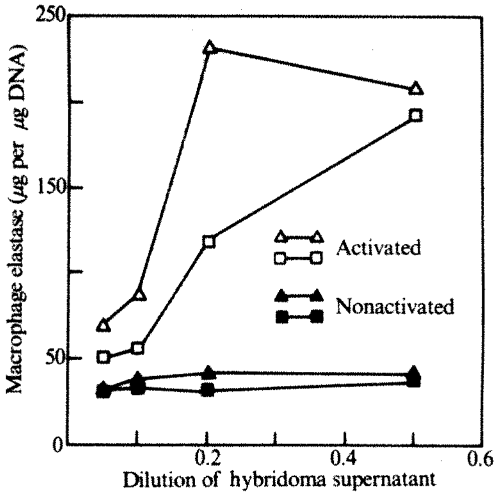


Fig. 1 Elastase production by macrophages activated with T-hybridoma W38C1 supernatants (activated, open symbols; nonactivated, solid symbols). Oil-induced C57BL/6 macrophages were prepared as for C2 production but after attachment and an overnight incubation with serial dilutions of hybridoma supernatant the monolayer was again washed three times and covered with 1 ml of Neuman–Tytell medium containing 60 μg ml^{-1} soybean trypsin inhibitor. After 60 h the medium was collected and dialysed overnight against 100 mM Tris-HCl *p*H 4.5 and lyophilized to dryness, and then reconstituted in one-tenth its original volume in 100 mM Tris *p*H 7.4 containing 5 mM $CaCl_2$ and assayed for elastase by the azocasein method[2]. The macrophage monolayer was lysed and DNA content measured by a fluorometric assay[9,10]. Results are expressed as μg equivalents of hog pancreatic elastase activity per μg DNA.

Table 1 shows the inhibition of migration induced by cloned hybridoma supernatants. In contrast to the colonies, all clones of W38 had migration indices of less than 1 but supernatants from clone C1 appeared even more active than supernatants from Con A-activated spleen cells. Results are given at the dilution of Con A supernatant which gave the greatest inhibition of migration and results were considered positive only if their migration index was ≤0.75 (ref. 6). The same supernatants were used to activate macrophages for production of C2. Table 1 shows that the same clone which was most active in the MIF assay gave a twofold enhancement of C2 production.

In two sets of experiments shown in Fig. 1, we measured elastase released into serum-free medium by a macrophage monolayer preincubated with serial dilutions of W38 C1 or control supernatants. In both experiments the supernatant from W38 C1 caused a three- to fourfold increase in production of elastase. DNA contents were not significantly different for either activated or non-activated macrophages, being 0.54 ± 0.098 and 0.545 ± 0.101 μg DNA per ml, respectively. W38 C1 supernatants have consistently been negative for interferon.

Our studies are the first to report a factor derived from a T hybridoma which directly affects macrophage function. Our data indicate that both migration inhibitory and macrophage activating activities are present in supernatants from the same cloned T-cell hybridoma. However, the results do not necessarily imply that they are the same molecule; indeed, multiple activities may be present and furthermore our factor(s) may be distinct from those which enhance macrophage cytotoxicity. These questions can be resolved only by chemical charac-

0028–0836/81/240502—02$01.00

terization of the active moieties present. We are now attempting to do this using macrophage elastase and plasminogen activator as markers for macrophage activation.

Our studies demonstrate that it is possible to preserve lymphokine production by activated spleen cells, by fusion to BW5147 whose supernatant has no biological activity in any of our assays. Establishment of a cell line offers several advantages: (1) production of a continuous level of activity relatively free from other mediators, (2) a homogeneous cell population for studying cell-surface markers, and (3) a source for extraction of specific mRNA.

We thank Dr Barbara Osborn for the gift of the BW5147 strain.

Received 19 December 1980; accepted 20 March 1981.

1. Mackaness, G. B. *J. exp. Med.* **129,** 973-992 (1969).
2. Werb, Z. & Gordon, S. *J. exp. Med.* **142,** 361-377 (1975).
3. Littman, B. H. & Ruddy, S., *J. exp. Med.* **145,** 1344-1352 (1977).
4. Remold, H. G. & David, J. R. *J. Immun.* **107,** 1090-1098 (1971).
5. Simpson, E. *et al. Curr. Topics Microbiol. Immun.* **81,** 195-201 (1978).
6. Nossal, R. in *Theoretical Immunology*, 121-142 (Dekker, New York, 1978).
7. McCoy, J. L., Dean, J. H. & Herberman, R. B. in *In Vitro Methods in Cell Mediated and Tumor Immunity* (eds Bloom, B. R. & David, J. R.) 621-628 (Academic, New York, 1976).
8. Borsos, T., Rapp, H. J. & Mayer, M. M. *J. Immun.* **87,** 310-324 (1961).
9. Kissane, J. M. & Robins, E. *J. biol. Chem.* **233,** 184-188 (1958).
10. Einstein, L. P., Schneeberger, E. E. & Colten, H. R. *J. exp. Med.* **143,** 114-126 (1976).

Synthesis of hepatitis B surface and core antigens in *E. coli*

Jeffrey C. Edman, Robert A. Hallewell*, Pablo Valenzuela, Howard M. Goodman & William J. Rutter

Department of Biochemistry and Biophysics, University of California, San Francisco, California 94143, USA

Hepatitis B virus (HBV) is the cause of a debilitating and potentially fatal disease, and possibly also of primary hepatocellular carcinoma[1]. A vaccine against hepatitis B would therefore be of considerable biomedical significance. One such vaccine, produced from the 22-nm particle form of hepatitis B surface antigen (HBsAg), has recently been found effective in clinical tests[2], but being derived from the sera of chronic carriers, it is expensive and its supply limited. The application of recombinant DNA methods can, in principle, provide a limitless source of vaccine. HBV DNA has been cloned in bacteria and the genes coding for HBsAg and hepatitis B core antigen (HBcAg) identified[3-5]. Small amounts of both antigens have been synthesized in *Escherichia coli*[3,6,7] and to enhance this synthesis we have constructed plasmids capable of expressing the genes for the antigens under the control of the efficient tryptophan (*trp*) operon regulatory region[8]. We describe here the construction of such recombinant plasmids which direct the synthesis of high levels of HBcAg and a β-lactamase: HBsAg fusion polypeptide.

The *trp* operon consists of a regulatory region and five coordinately expressed structural genes, *trp*E to *trp*A, respectively (see Fig. 1*a* and ref. 9). The leader peptide of the *trp* operon (*trp*L)[10] has a *Taq*I site between the Shine–Dalgarno sequence (SD)[11] and the initiator AUG codon (Fig. 1*b*). The location of this site facilitates the construction of a vector suitable for the creation of hybrid ribosome-binding sites, like those described previously for vectors based on the *lac* operon[12,13], containing the *trp*L SD sequence and the initiation codon of the foreign gene. Although the synthesis of the 14-amino acid *trp* leader peptide has not been demonstrated, an *in vivo* recombinant between the *trp*L SD sequence and the initiation codon of *trp*C has been shown to direct efficient *trp*C synthesis[14]. In addition, some *trp* operon deletion mutants between *trp*L and *trp*E have been shown to produce *trp*L : *trp*E fusion polypeptides, indicating that *trp*L does contain a functional ribosome-binding site[15].

Plasmid ptrpL1 was constructed from plasmid ptrpE2-1 (see ref. 16 and Fig. 1*c*) by first isolating the 32-base pair *Hpa*I–*Taq*I fragment of ptrpE2-1 by polyacrylamide gel electrophoresis (PAGE). This fragment contains part of the *trp* promoter-operator and the SD sequence of *trp*L[10]. The *Hpa*I–*Taq*I fragment was recloned into *Hpa*I–*Cla*I-digested ptrpE2-1. Because the *Taq*I sequence (T/CGA) of the *Hpa*I–*Taq*I fragment is preceded by an A residue on its 5′ side, this construction resulted in the formation of a new *Cla*I site (AT/CGAT). The sequence of this junction is shown in Fig. 1*d*. The resulting plasmid, ptrpL1, has a unique *Cla*I site, three base pairs 3′ to the SD sequence of trpL, which is useful because several different restriction fragments with 5′ CG single-stranded termini can be cloned into it (for example, *Hpa*II, *Taq*I, *Acy*I).

Figure 2 outlines the procedure used to construct a series of plasmids that direct the synthesis of HBcAg sequences. A 1,005-base pair *Hha*I fragment of HBV DNA[5] containing all the HBcAg gene was purified by PAGE and treated with *Hpa*II methylase to prevent cleavage of the fragment during linker digestion with *Hpa*II. Our HBV DNA clone contains two potential initiation codons for HBcAg[5]. The start site for HBcAg is not known, but most of the evidence is consistent with translation initiation at the second codon[3]. This work utilizes the first initiation codon which should result in the synthesis of a polypeptide containing the entire amino acid sequence for HBcAg and probably an additional 29 amino acids at the N terminus. The initiation codon for HBcAg lies 15 base pairs from one end of this fragment. This fragment was treated with T4 DNA polymerase in the absence of deoxynucleoside triphosphates (to allow the 3′ exonuclease to function), in conditions where 0–10 nucleotides are removed from each 3′ end, then with S_1 nuclease. *Bam*HI linkers (CCGGATCCGG, which also contain the *Hpa*II site, CCGG) were added and the fragments digested with *Hpa*II to generate the 5′CG cohesive end. After purification, these fragments were ligated with *Cla*I-digested and phosphatase-treated ptrpL1. The ligation mixture was used to transform HB101 and recombinants were screened for HBcAg by radioimmune assay (RIA). Seventeen out of 40 recombinants tested were HBcAg positive. Restriction enzyme analysis of the plasmids from HBcAg-positive clones showed that all contained HBcAg sequences and were in the proper orientation for *trp*-dependent expression of HBcAg. Plasmids with the HBcAg gene in the opposite orientation were HBcAg negative.

DNA sequence analysis of four positive clones confirmed that the initiation codon of HBcAg was in close proximity (13–16 base pairs) to the SD sequence of the leader peptide. An HBcAg standard was not available to measure the absolute levels of HBcAg produced in *E. coli*, but the relative levels varied fourfold. The plasmid responsible for the highest level of expression, pCA246, whose sequence is presented at the bottom of Fig. 2, was chosen for further characterization. The amount of HBcAg produced in a 20-min labelling period is 3.2% of the newly synthesized protein uncorrected for cysteine content (~150,000 molecules per cell) as estimated by densitometric scanning of the autoradiogram, see Fig. 3*A*, lane *d*. The level of HBcAg can be increased to 10% of newly synthesized protein by optimizing the concentration of the inducer, 3-*β* indolylacrylic acid (IA, unpublished results).

The efficacy of the *trp* promoter in enhancing expression of a linked coding sequence has been demonstrated by the significant production of a fused protein involving the *trp*D gene and human growth hormone coding sequences[17], and by the overproduction of a contiguous *β*-lactamase gene (Fig. 3*A*, lane *b*), which constitutes 12% of labelled protein in cells containing ptrpL1, as a result of partial derepression of *trp* promoters in cells grown in media lacking tryptophan. On addition of IA all *trp* promoters become fully derepressed[8,9] and the level of *β*-lactamase increases to 50% of labelled protein (Fig. 3*A*, lane *e*).

---

* To whom enquiries concerning the *trp* vectors should be addressed.

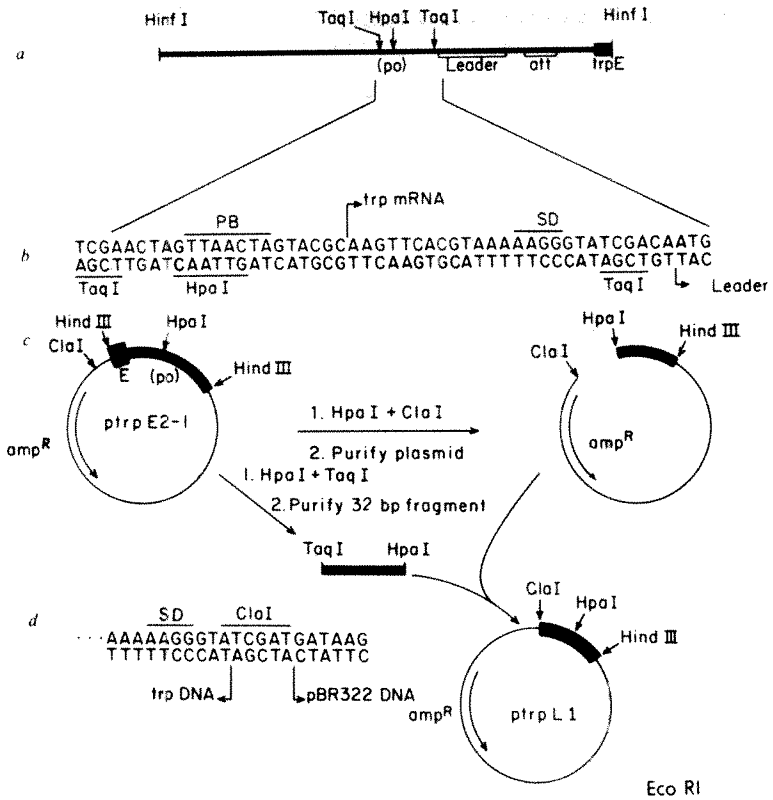


Fig. 1 Construction of expression vector ptrpL1. *a*, Partial restriction map of the 492-base pair *Hin*fI fragment of the *trp* operon of *E. coli*. The promoter–operator (po), leader region, attenuator (att) and *trp*E structural gene are shown in their approximate locations[10]. *b*, DNA sequence[24] of the promoter–operator and proximal portion of the leader region. Note that a *Taq*I site lies between the Shine–Dalgarno sequence (SD) and the initiation codon (ATG) of the leader region. The proposed Pribnow box (PB) and transcription initiation site are also shown. *c*, Construction of ptrpL1. Plasmid vector ptrpE2-1 (construction to be published elsewhere, see text) was digested with *Hpa*I and *Taq*I and the 32-base pair fragment of the *trp* operon purified by PAGE[25,26]. Separately, ptrpE2-1 was digested with *Cla*I and *Hpa*I and the plasmid purfied by agarose gel electrophoresis[26]. The plasmid vector (200 ng) was ligated to the 32-base pair fragment (5 ng) in 20 μl with T4 DNA ligase and the ligation mix used to transform HB101 (ref. 27). Ampicillin-resistant transformants were analysed for plasmids containing single *Cla*I and *Hind*III sites[28]. Several clones were analysed by restriction mapping and one by DNA sequence[29]. This plasmid had the expected sequence around the *Cla*I site and was designated ptrpL1. *d*, DNA sequence around the *Cla*I site of ptrpL1 showing the location of the *trp*L SD sequence. Enzymes were purchased from either New England Biolabs or Boehringer-Mannheim and reactions carried out according to manufacturer's instructions.

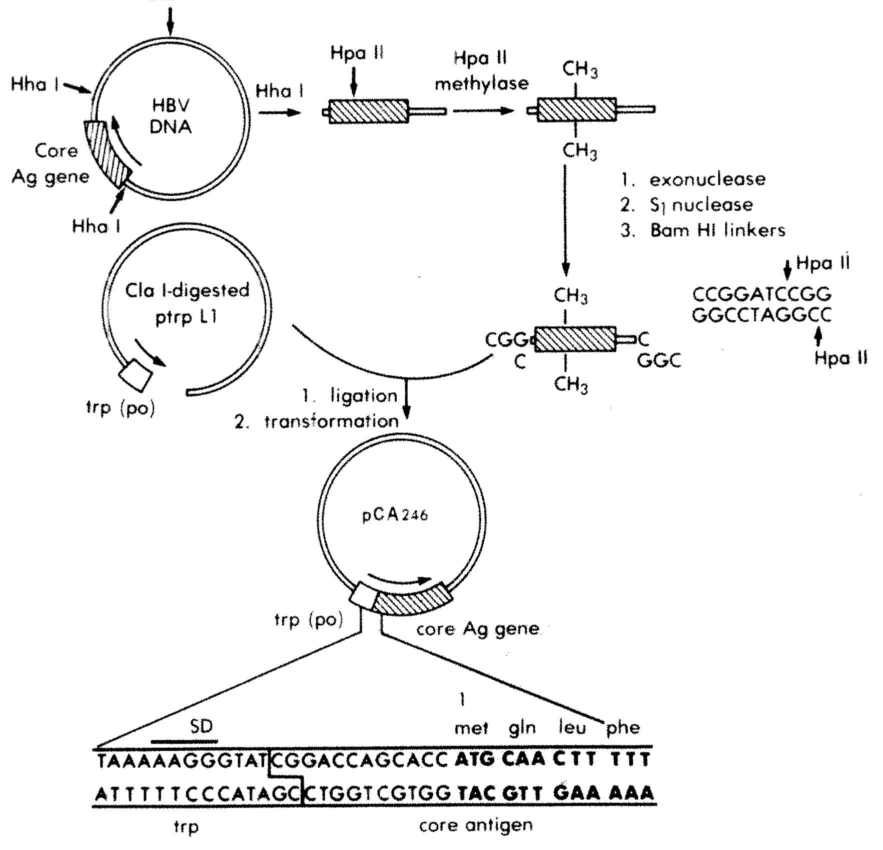


Fig. 2 Construction of plasmids expressing HBcAg. Cloned HBV DNA[5] was digested with *Hha*I and a 1,005-base pair fragment containing the HBcAg gene was isolated by PAGE[25,26]; 20 μg of this fragment were then treated with *Hpa*II methylase[30] (a gift from K. Agarwal). After phenol–chloroform extraction and ethanol precipitation the fragment was treated with 28 U of T4 DNA polymerase (prepared by J. Barry) in the absence of dNTPs (to allow the 3′ exonuclease activity to function) in 200 μl of 30mM Tris-acetate, *p*H 7.8, 67 mM KOAc, 10 mM MgOAc, 0.5 mM dithiothreitol and 100 μg ml^{-1} bovine serum albumin for 30 s at 37 °C. The reaction was stopped with phenol, extracted with chloroform and precipitated with ethanol. The fragments were then treated with S$_1$ nuclease and *Bam*HI linkers (Collaborative Research) added using T4 DNA ligase[31]. The fragments were digested with *Hpa*II and purified from the digested linkers by PAGE and 150 ng were ligated with 200 μg *Cla*I-digested and calf intestine alkaline phosphatase-treated ptrpL1 (10 U sigma type VII phosphatase incubated with 20 μg DNA at 68 °C for 30 min in 400 μl 10 mM Tris, *p*H 7.5, 0.5 mM EDTA) in 20 μl and the ligation mixture used to transform HB101 (ref. 32). Transformants were selected on L plates[33] containing 20 μg ml^{-1} ampicillin and screened for recombinant plasmid DNA[34]. Forty such recombinants were then tested for HBcAg using a double antibody RIA with human anti-HBcAg IgG[35]; 17 were HBcAg positive. Analysis of DNA after digestion with restriction enzymes showed that all 17 contained HBcAg DNA sequences in a *trp*-dependent expression orientation. Four were chosen for DNA sequence analysis. The sequence of the plasmid producing the highest levels of HBcAg, pCA246, is shown. The sequence of the others was slightly longer having one or three additional nucleotides of HBV DNA.

When cells containing pCA246 were treated with IA[9], labelled with ^{35}S-cysteine and the products run on SDS gels, an increased synthesis of a 22,000-molecular weight polypeptide not present in the control strain was observed (Fig. 3*A*, lanes *a*, *d*). The DNA sequence of the HBcAg gene predicts a protein of this molecular weight[5]. In such cells there is also a strong polar effect on β-lactamase production; the decreased level of β-lactamase may be due to a transcription termination site in the HBV DNA or inefficient transcription/translation of HBcAg sequences.

Anti-HBcAg serum precipitates (Fig. 3*B*, lanes *a*, *b*) a 22,000-MW polypeptide, corresponding to the HBcAg protein predicted by the constructed coding sequence, and a polypeptide of MW ~19,000 that is not evident in the labelled proteins resolved on SDS gels (Fig. 3*A*, lanes *a*, *d*). This latter polypeptide probably results from initiation of translation at the second methionine of the 22,000-MW polypeptide or degradation. The pattern of background bands precipitated with anti-HBcAg is different from that in the normal serum control. Sedimentation studies of HBcAg produced in *E. coli* demonstrate a sedimentation coefficient >100S, suggesting that HBcAg is present as some tightly associated complex. As HBcAg functions as a DNA-binding protein[3], it may be immunoprecipitated as a complex with DNA or RNA. This would result in the precipitation of other proteins bound to nucleic acids and explain the character of the immunoprecipitates.

It has been suggested that the differential utilization of codons for particular amino acids between *E. coli* and certain eukaryotes[18] may account for the observation that some eukaryotic cDNAs are expressed at lower levels than expected in *E. coli*[17,19]. One mechanism proposed for this effect is that cognate tRNA molecules are in limiting concentrations for rarely used codons, thus preventing high-level expression of a gene containing several such codons. The HBcAg gene contains 25 arginine codons, 17 of which are rarely utilized in *E. coli*[5,18]. We have shown that HBcAg is produced in substantial quantities when placed under the auspices of ptrpL1. Thus at least in this instance, inclusion of many rarely utilized codons does not prevent high-level expression of a gene sequence in *E. coli.*

Plasmids similar to the HBcAg-expressing ones were constructed using the HBsAg gene. Cells transformed with these plasmids were uniformly negative for HBsAg by RIA. Several clones were analysed by restriction mapping and two by DNA sequence. The distance separating the SD sequence and the initiation codon for HBsAg was 11 and 12 nucleotides. Neither of the clones accumulated polypeptides of the proper size (MW 25,398) on addition of IA, but the polar effect on β-lactamase was still present (data not shown). This result could have been due to rapid degradation or inefficient transcription/translation.

HBsAg contains a prominent hydrophobic region as well as an N-terminal region resembling a modified signal peptide. There are several studies of powerful peptide degradation systems in bacteria including an ATP-dependent system[20] probably operative at the cell membrane. We reasoned that elimination of part of the sequence coding for the N-terminal region and subsequent replacement with a nucleotide sequence coding for a stable protein fragment might enhance the stability of this molecule. Because β-lactamase is overproduced in cells containing ptrpL1, we elected to try to produce a β-lactamase: HBsAg fusion polypeptide. Figure 4 summarizes the construction of such plasmids. *Pst*I-digested ptrpL1 was 'tailed' with deoxyguanosine using terminal transferase in conditions where ~15 dG residues were added to each 3′ end. A 744-base pair *Hinc*II fragment of HBV DNA containing the coding sequence for 204 amino acids (the amino terminal 22 are missing) of HBsAg was purified by PAGE, tailed with dC and hybridized with the tailed ptrpL1. The mixture was used to transform *E. coli* HB101.

The DNA sequence of this HBV DNA fragment predicts that when the HBsAg sequences are in the proper orientation and same phase as β-lactamase, a 43,000-MW polypeptide should be produced. This includes 183 amino acids of pre-β-lactamase (MW 20,100), 5–10 glycine residues from the G tail (MW 375–750), and 204 amino acids of HBsAg (22,600). If the pre-sequence is cleaved, the fusion protein would have a predicted MW of 41,000. Thirty-eight ampicillin-sensitive colonies were grown in the presence of IA, labelled, and the products run on SDS gels. Two clones producing a 41,000-MW polypeptide as a major product (Fig. 3*A*, lanes *c*, *f*) were chosen for DNA sequence. As expected, the HBsAg sequences were in the same phase as the amino-terminal β-lactamase sequences (Fig. 4). This fusion protein represents 8.5% of protein (uncorrected for cysteine content) synthesized in a 20-min period (~170,000 molecules per cell) as estimated by densitometric scanning of the autoradiogram.

The presumptive 41,000-MW fusion polypeptide is immunoprecipitated with anti-HBsAg IgG (Fig. 3*B*, lane *e*). Addition of cold HBsAg inhibits the precipitation completely (Fig. 3*B*, lane *d*). Thus, HBsAg determinants are present in the fusion polypeptide. The level of the 41,000-MW polypeptide varied widely in different experiments and in some cases was not detected by

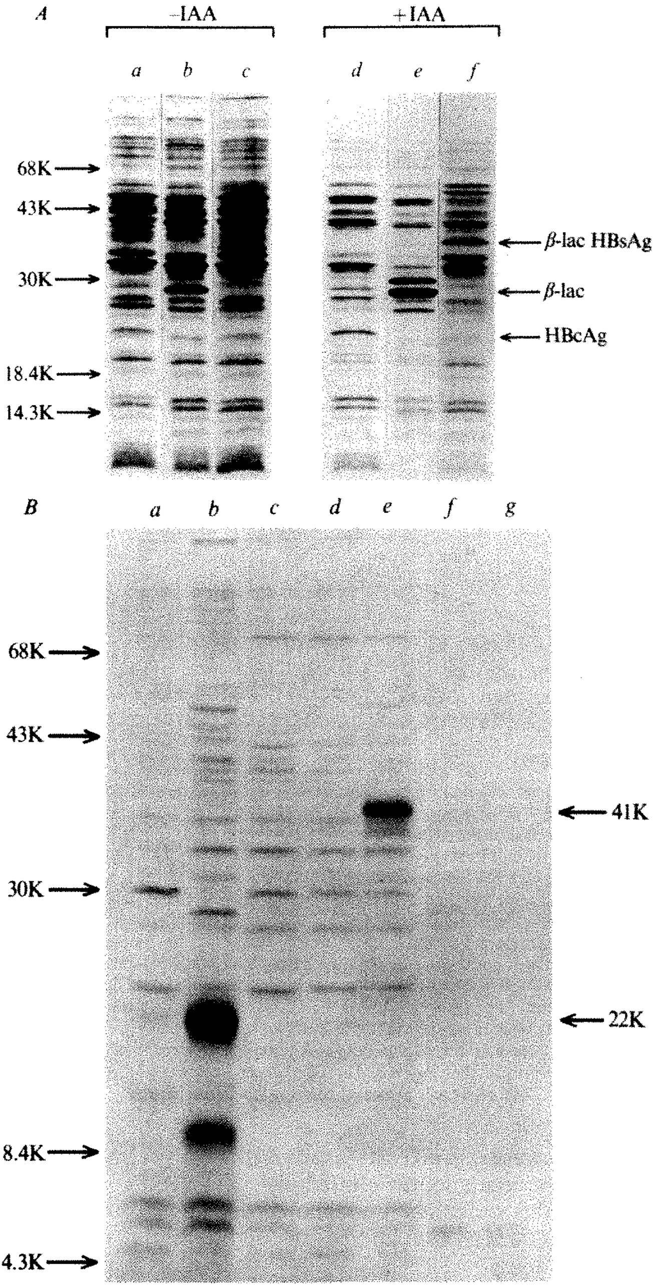


Fig. 3 *A*, autoradiography of SDS-polyacrylamide gel[36] of proteins produced in cells containing pCA246 and pSA4A. Overnight cultures were grown in M9 medium[33] containing 0.25% casamino acids, 0.5% glucose and 0.01% thiamine. The cultures were diluted 1 : 10 with fresh media, grown for 1 h at 30 °C, 15 μg ml⁻¹ of IA were added when induction of the *trp* operon was required, and cultures grown for another 2 h at 30 °C (ref. 8). The cultures (1 ml) were then labelled for 20 min at 30 °C with 10 μCi ml⁻¹ of ^{35}S-cysteine (NEN), labelled cells collected in 1.5-ml Eppendorf tubes and resuspended by boiling for 2 min in Laemmli sample buffer[36]. Samples were run on 12.5% Laemmli gels[36], the gels were dried and subjected to autoradiography. Lanes *a*–*c* are from uninduced cells containing pCA246 (HBcAg), ptrpL1 and pSA4A (β-lactamase: HBsAg fusion), respectively. Lanes *d*–*f* are samples induced with IA of cells containing pCA246, ptrpL1 and pSA4A, respectively. Molecular weight markers on the left: bovine serum albumin (68,000), ovalbumin (43,000), carbonic anhydrase (30,000), β-lactoglobulin (18,400) and lysozyme (14,300). Arrows on right of the bands denote the migration of the β-lactamase : HBsAg fusion (MW 41,000), β-lactamase (29,000) and HBcAg from pCA246 (22,000). *B*, autoradiogram of SDS-polyacrylamide gel of immunoprecipitates of HBcAg and β-lactamase : HBsAg fusion produced in *E. coli*. ^{35}S-cysteine-labelled samples were prepared as described above, except labelled cells were resuspended in phosphate-buffered saline containing 1 mM phenylmethylsulphonylfluoride. Cells were sonicated and proteins immunoprecipitated using the SAC technique[17]. *a*, pCA246 + normal human IgG; *b*, pCA246 + human anti-HBcAg; *c*, pSA4A + normal guinea pig IgG; *d*, pSA4A + guinea pig anti-HBsAg and 2 μg unlabelled HBsAg; *e*, pSA4A + guinea pig anti-HBsAg; *f*, ptrpL1 + anti-HBcAg; and *g*, ptrpL1 + anti-HBsAg. One μg of each antibody (Merck) was added per tube. Markers on left as above. Numbers on right refer to the β-lactamase : HBsAg fusion (MW 41,000) and HBcAg (22,000) respectively.

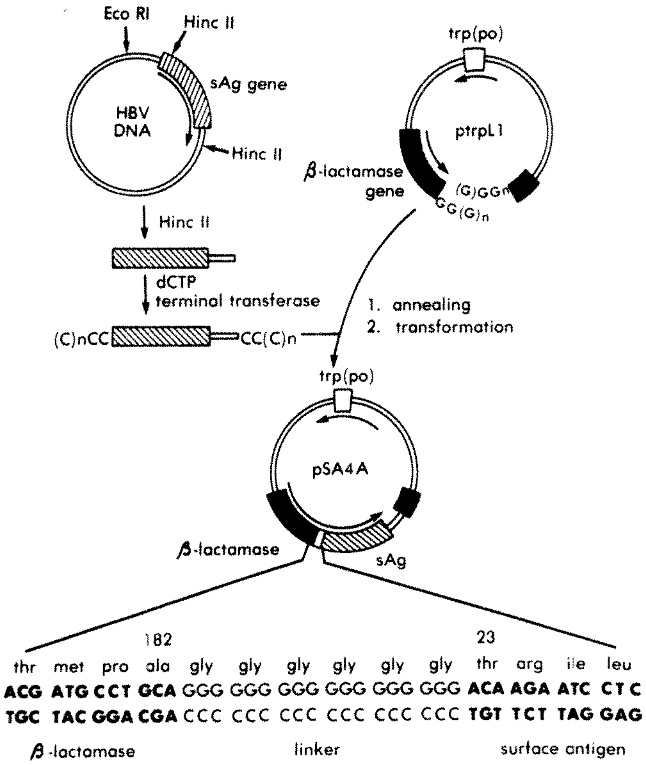

Eco RI
Hinc II
sAg gene
HBV DNA
Hinc II
Hinc II
dCTP terminal transferase
(C)nCC
CC(C)n
trp(po)
ptrpL1
β-lactamase gene
(G)GG
G(G)n
1. annealing
2. transformation
trp(po)
pSA4A
β-lactamase
sAg
182
23
thr met pro ala gly gly gly gly gly gly thr arg ile leu
ACG ATG CCT GCA GGG GGG GGG GGG GGG GGG ACA AGA ATC CTC
TGC TAC GGA CGA CCC CCC CCC CCC CCC CCC TGT TCT TAG GAG
β-lactamase
linker
surface antigen

Fig. 4 Construction of plasmids expressing a β-lactamase : HBsAg fusion. Cloned HBV DNA was digested with *Hin*cII and a 744-base pair fragment containing part of the HBsAg gene was isolated by PAGE. This fragment was tailed with approximately 15 dC residues using terminal transferase[37]. Plasmid ptrpL1 was digested with *Pst*I and tailed with ~15 dG residues. The fragment (35 ng) was then hybridized with the tailed plasmid (200 ng)[38] and the mixture used to transform HB101. Transformants were selected on L plates containing 2.5 μg ml^{-1} tetracycline[31] and tested for ampicillin resistance on plates containing 20 μg ml^{-1} ampicillin. Samples were prepared from 38 ampicillin-sensitive transformants, induced with IA as described in Fig. 3 legend. Samples were run on 12.5% Laemmli gels[36], the gels were dried and subjected to autoradiography. Two clones, pSA4A and pSA7A, producing large amounts of a 41,000-MW polypeptide were found and these were chosen for further analysis. The DNA sequence of the fusion junction of pSA4A is shown at the bottom of the figure.

PAGE. However, in all cases a radioactive polypeptide of MW 41,000 was immunoprecipitated by anti-HBsAg IgG. The reason for this variability is not known, but may be due to either subtle variations in cellular growth conditions and/or shifting product stability.

We have presented here evidence that substantial quantities of hepatitis B surface and core antigen sequences can be synthesized in *E. coli*. These proteins may be useful in the development of an HBV vaccine. HBsAg is synthesized as a fusion polypeptide that contains an antigenic determinant found in native HBsAg; its effectiveness as an immunogen in generating protective antibodies is under investigation. The product will probably be less effective than the HBsAg particle itself because it has been demonstrated that dissociated HBsAg is much less immunogenic than the 22-nm HBsAg particle[21]. However, this disadvantage may not be decisive as HBsAg can be produced in bacteria in any desired quantity.

The availability of HBcAg synthesized in bacteria may answer some questions about the biology of HBV. For example, it has been suggested[22,23] that the cryptic 'e' antigen of HBV is a dissociated form of HBcAg protein, but the preparations used may have been contaminated by other antigens. As the HBcAg produced in bacteria has no other HBV gene products, this possibility can now be tested directly.

It is also possible through similar recombinant DNA methods to identify and express the remaining genes of HBV (putative e antigen, DNA polymerase) and evaluate their role in the cytopathology of HBV. These studies might elucidate the aetiology of the disease and suggest new strategies for its control.

Received 8 December 1980; accepted 20 March 1981.

1. Zuckerman, A. J. *J. Tox. envir. Hlth* **5,** 275–280 (1979).
2. Szmuness, W. *et al. New Engl. J. Med.* **303,** 833–841 (1980).
3. Pasek, M. *et al. Nature* **282,** 575–579 (1979).
4. Galibert, F. *et al. Nature* **281,** 646–650 (1979).
5. Valenzuela, P. *et al.* in *Animal Virus Genetics* (ed. Fields, B. N., Jaenisch, R. & Fox, C. F.) 57–71 (Academic, New York, 1980).
6. Burrel, C. J. *et al. Nature* **279,** 43–47 (1979).
7. Charnay, P. *et al. Nature* **286,** 893–895 (1980).
8. Hallewell, R. A. & Emtage, S. *Gene* **9,** 27–47 (1980).
9. Morse, D. E., Mosteller, R. D. & Yanofsky, C. *Cold Spring Harb. Symp. quant. Biol.* **34,** 725–740 (1969).
10. Lee, F., Bertrand, K., Bennett, G. & Yanofsky, C. *J. molec. Biol.* **121,** 193–227 (1978).
11. Shine, J. & Dalgarno, L. *Nature* **254,** 34–38 (1975).
12. Goeddel, D. *et al. Nature* **281,** 544–548 (1979).
13. Roberts, T. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **76,** 5596–5600 (1979).
14. Christie, G. & Platt, T. *J. molec. Biol.* **143,** 335–341 (1980).
15. Miozzari, G. F. & Yanofsky, C. *J. Bact.* **133,** 1457–1466 (1978).
16. Tacon, W. *et al. Molec. gen. Genet.* **177,** 427–438 (1980).
17. Martial, J. A., Hallewell, R. A., Baxter, J. D. & Goodman, H. M. *Science* **205,** 602–607 (1979).
18. Grantham, R., Gautier, C., Gouy, M., Mercier, R. & Pavez, A. *Nucleic Acids Res.* **8,** 49–61 (1980).
19. Fraser, T. H. & Bruce, B. J. *Proc. natn. Acad. Sci. U.S.A.* **75,** 5936–5940 (1978).
20. Murakami, K., Voellmy, K. & Goldberg, A. L. *J. biol. Chem.* **254,** 8194–8200 (1979).
21. Cabral, G. A. *et al. J. gen. Virol.* **38,** 339–350 (1978).
22. Takahashi, K. *et al. J. Immun.* **122,** 275–279 (1979).
23. Ohori, H. *et al. Intervirology* **13,** 74–82 (1980).
24. Bennett, G. N., Schweingruber, M. E., Brown, K. D., Squires, C. & Yanofsky, C. *Proc. natn. Acad. Sci. U.S.A.* **73,** 2351–2355 (1976).
25. Maniatis, T., Jeffrey, A. & van de Sande, H. *Biochemistry* **14,** 3787–3794 (1975).
26. Smith, H. O. *Meth. Enzym.* **65,** 371–380 (1980).
27. Boyer, H. W. & Roulland-Dussoix, D. *J. molec. Biol.* **41,** 459–472 (1969).
28. Birnboim, H. C. & Doly, J. *Nucleic Acids Res.* **7,** 1513–1523 (1979).
29. Maxam, A. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74,** 560–564 (1977).
30. Yoo, J. O. & Agarwal, K. L. *J. biol. Chem.* **255,** 6445–6449 (1980).
31. Ullrich, A. *et al. Science* **196,** 1313–1319 (1977).
32. Cohen, S. N., Chang, A. C. Y. & Hsu, L. *Proc. natn. Acad. Sci. U.S.A.* **69,** 2110–2114 (1972).
33. Miller, J. H. in *Experiments in Molecular Genetics*, Appendix I (Cold Spring Harbor Laboratory, New York, 1972).
34. Barnes, W. M. *Science* **195,** 393–395 (1977).
35. Ling, C. M. & Overby, L. R. *J. Immun.* **109,** 834–841 (1972).
36. Laemmli, U.K. *Nature* **227,** 680–685 (1977).
37. Roychoudhury, R., Jay, E. & Wu, R. *Nucleic Acids Res.* **3,** 863–877 (1976).
38. Chang, A. C. Y. *et al. Nature* **275,** 617–624 (1978).

Lack of actin III in fibrillar flight muscle of flightless *Drosophila* mutant *raised*

Alois B. Lang, Christoph Wyss & Hans M. Eppenberger*

Institute for Cell Biology, Swiss Federal Institute of Technology, Hönggerberg, CH-8093 Zürich, Switzerland

Actin heterogeneity is well known: within one vertebrate species up to six forms of actin, differing in isoelectric point, amino acid composition and antigenic sites have been described and their synthesis shown to be tissue specific[1–8]. In *Drosophila melanogaster* actin genes have been located at six distinct sites on the chromosomes[9,10] and three forms of actin (synthesized and/or accumulated in different amounts in different tissues[11–14]) have been distinguished electrophoretically. In view of the obvious importance of actin in muscle, it was of interest to search in *Drosophila* for abnormalities of actin in flightless mutants which primarily affect fibrillar flight muscles, such as *raised* (*rsd*)[14–20]. Here we report that the actin pattern in two types of adult skeletal muscles is differently affected by the *rsd* mutation. Since *rsd* is located on the right arm of the third chromosome[25], which is not the known position of the structural genes of actin[9,10], we conclude that *rsd* is a regulatory gene.

In tubular, synchronous muscle (leg and direct flight muscles) actin I is synthesized in addition to the (cytoplasmic) actins II and III which are found also in non-muscle tissue. In fibrillar, asynchronous (indirect flight) muscle, however, no synthesis of actin I is detectable. Actin III, which in other tissues appears to be short lived and detectable only in minor quantities, is accumulated in fibrillar muscle to levels comparable with those of actin II. Analysis of the actin pattern in the mutant *rsd* reveals no synthesis of actin III in fibrillar muscle whereas tubular muscle shows the same pattern as in wild-type flies. At the ultrastructural level fibrillar muscles of *rsd* lack thin filaments whereas tubular muscles appear normal. This suggests that the stable actin III found in the wild type represents a form of actin specific for fibrillar muscle.

---

* To whom reprint requests should be addressed.

To detect specific effects of the mutation we analysed tissues with as little contamination by other tissue types as possible[15,16]. For tubular muscle a preparation of legs with musculature attached—in particular the largest tubular muscle, the tergal depressor of the trochanter—proved to be satisfactory both for handling and for muscle homogeneity. A preparation of the thorax without legs and emptied of internal organs was taken to represent fibrillar musculature (dorsolongitudinal and dorsoventral muscles attached to the cuticle). After *in vitro* labelling these samples were analysed for protein synthesis (by fluorography) and accumulation pattern (Serva blue stain) on two-dimensional gel electrophoresis (Figs 1, 2).

Figure 1 shows that in fibrillar muscle of wild-type flies no actin I is made whereas actin III (and II, which seems to be present in all cell types) is synthesized and accumulated. In contrast, in tubular muscle all three actins are synthesized and accumulated. The accumulation of actin III is less prominent than in fibrillar muscle. Although the data on fibrillar muscle largely confirm the findings of Horovitch *et al.*[11], the presence of actin I in tubular muscle contradicts their hypothesis that actin I, detected so far in larval body wall muscles, larval and adult visceral and adult abdominal muscles, and embryonic myogenic cultures[11-13], might be characteristic of supercontracting muscle.

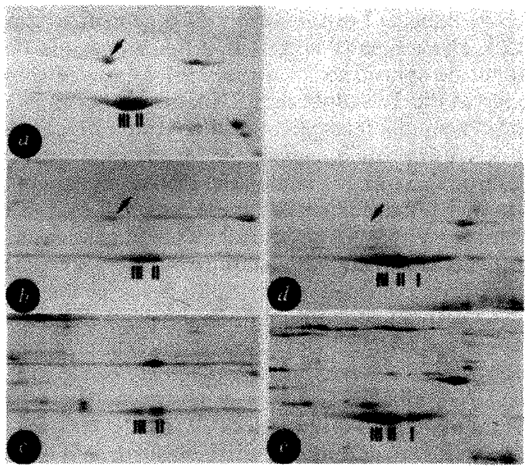


Fig. 1 Analysis of actin accumulation (*a*, *b*, *d*; Serva blue stain) and synthesis (*c*, *e*; by fluorography) in adult fibrillar flight muscles (*a*–*c*) and adult tubular muscles (*d*, *e*) of *Drosophila melanogaster* (Oregon-R). Flies were immobilized on ice and immediately dissected. The samples used were: fibrillar muscles attached to the cuticle of the thorax and tubular muscle of the tergal depressor of the trochanter still attached to the second leg (leg muscles). These were prepared in modified ZH medium[21], but lacking methionine, yeast extract components and horse serum. For labelling they were transferred to the same medium supplemented with 3 mCi ml^{-1} ^{35}S-methionine (NEN, 1,108.5 Ci mmol^{-1}) and incubated for 90 min at 26 °C. Sample preparation, two-dimensional polyacrylamide gel electrophoresis (isoelectrofocusing between *p*H 5 and 8 in the first dimension and 14% SDS-gel electrophoresis in the second dimension), Serva blue staining of the gel and fluorography were performed using published procedures[2,22,23]. Total protein extracts of fibrillar flight muscles of 5–10 h (*a*) and 5-day-old flies (*b*, *c*) and total protein extracts of tubular muscles of 5-day-old flies (*d*, *e*) were separated. For fibrillar and for tubular muscle the samples applied to the gels correspond to the material from 3 and 6 flies respectively. The position of actins I, II and III and of an unknown protein of ~52,000 molecular weight (arrow) are indicated. The actin region of the gels are presented with the more acidic polypeptides to the left.

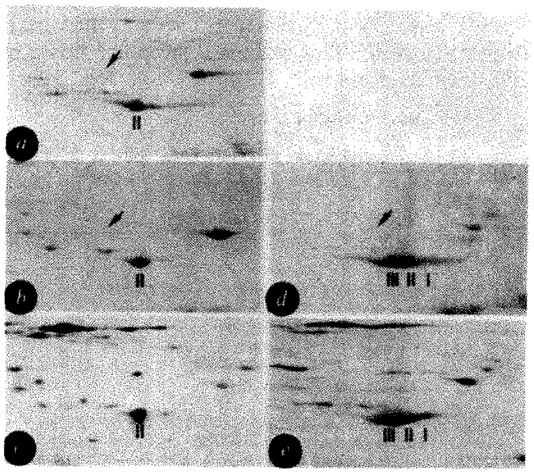


Fig. 2 Analysis of actin accumulation (*a*, *b*, *d*; staining) and synthesis (*c*, *e*; fluorography) in adult fibrillar flight muscles (*a*–*c*) and adult tubular muscles (*d*, *e*) of the mutation raised (*rsd*). Sample preparation and processing and symbols as for Fig. 1. For fibrillar and for tubular muscle the samples applied to the gels correspond to the material from three and six flies respectively. Samples are from newly emerged flies (*a*) and from 5-day-old flies (*b*–*d*).

The protein pattern of tubular muscle of *rsd* flies is indistinguishable from that of wild-type flies in both accumulation and synthesis (Fig. 2). When, however, the pattern in fibrillar muscle is analysed, important differences are found. Most obvious is the fact that in fibrillar muscle of *rsd* only actin II can be detected. Because ultrastructure studies show that fibrillar muscle of *rsd* lacks thin filaments (C. A. Reinhardt and A. B. Lang, unpublished data) it may be that in fibrillar muscle thin filaments are normally composed of actin III whereas those of tubular muscle represent actin I. Note that actin III is not even detected as the short-lived component thought to be ubiquitous. If this is so, it must be concluded that the precursor product relationship of actins III and II indicated by the data of Berger and Cox in Kc cells[14] does not hold for fibrillar muscle. It remains an open question whether the electrophoretically resolved actin form III contains more than one protein species—one accumulated in normal fibrillar muscle, another rapidly metabolized in other cell types. These 'two forms' of actin III could conceivably be products of different actin genes, one of them affected by *rsd*.

Some flightless mutants like *heldup-2* (*wingsup A*)[17-20] or *indented thorax*[19] are known to cause degeneration of flight muscle, so a comparison of two developmental stages (newly emerged and 5-day-old flies) was undertaken. Examination of Fig. 2*a*, *b* reveals identical protein pattern for 1-day and 5-day-old fibrillar flight muscle of *rsd*. This indicates that *rsd* does not cause a premature degeneration.

A further protein, of molecular weight 52,000, is also apparently absent from extracts of *rsd* fibrillar muscle. This protein seems to be a specific component of fibrillar muscle because it is also absent in wild-type tubular muscle. Nothing is known about the function of this protein, although it might be identical with a similar, apparently actin-associated component found in muscle of *Lethocerus*[24].

Actin gene sequences have been identified by *in situ* hybridization at six sites on the *Drosophila* chromosome[9,10]. Because none of them corresponds to the genetically mapped site of *rsd* (3–95.4; P. T. Ives cited in ref. 25) and tubular muscle is not affected, we conclude that *rsd* is not an actin gene mutation but rather a tissue-specific regulatory element affecting actin III metabolism. As several other proteins are affected in addition to actin III, it is not clear at what level this regulation occurs.

We are indebted to the late Professor I. I. Deak and his co-workers (University of Zürich) for supplying mutants and for stimulating discussions. This work was supported by a predoctoral ETH-training grant to A.B.L.

Received 21 January; accepted 16 April 1981.

1. Whalen, R. G., Butler-Browne, G. S. & Gros, F. *Proc. natn. Acad. Sci. U.S.A.* **73,** 2018–2022 (1976).
2. Garrels, J. I. & Gibson, W. *Cell* **9,** 793–805 (1976).
3. Zechel, K. & Weber, K. *Eur. J. Biochem.* **89,** 105–112 (1978).
4. Elzinga, M. & Lu, R. C. in *Contractile Systems in Non-Muscle Tissue* (eds Perry, S. V., Margreth, A. & Adelstein, R. S.) 29–37 (North-Holland, Amsterdam, 1976).
5. Vandekerckhove, J. & Weber, K. *Differentiation* **14,** 123–133 (1979).
6. Vandekerckhove, J. & Weber, K. *FEBS Lett.* **102,** 219–222 (1979).
7. Morgan, J. L., Holladay, C. R. & Spooner, B. S. *Proc. natn. Acad. Sci. U.S.A.* **77,** 2069–2073 (1980).
8. Vandekerckhove, J. & Weber, K. *J. molec. Biol.* **126,** 783–802 (1978).
9. Fyrberg, E. A., Kindle, K. L. & Davidson, N. *Cell* **19,** 365–378 (1980).
10. Tobin, S. L., Zulauf, E., Sanchez, F., Craig, E. A. & McCarthy, B. J. *Cell* **19,** 121–131 (1980).
11. Horovitch, S. J., Storti, R. V., Rich, A. & Pardue, M. L. *J. Cell Biol.* **82,** 86–92 (1979).
12. Fyrberg, E. A. & Donady, J. J. *Devl Biol.* **68,** 487–502 (1979).
13. Storti, R. V., Horovitch, S. J., Scott, M. P., Rich, A. & Pardue, M. L. *Cell* **13,** 589–598 (1978).
14. Berger, E. & Cox, G. *J. Cell Biol.* **81,** 680–683 (1979).
15. Lang, A. B., Wyss, C. & Eppenberger, H. M. *J. Muscle Res. Cell Motility* **1,** 147–161 (1980).
16. Lang, A. B., Wyss, C. & Eppenberger, H. M. *J. Muscle Res. Cell Motility* **1,** 478 (1980).
17. Hotta, Y. & Benzer, S. *Nature* **240,** 527–535 (1972).
18. Benzer, S. *Scient. Am.* **229,** 24–37 (1973).
19. Deak, I. I. *J. Embryol. exp. Morph.* **40,** 35–63 (1977).
20. Koana, T. & Hotta, Y. *J. Embryol. exp. Morph.* **45,** 123–143 (1978).
21. Wyss, C. *Somat. Cell Genet.* **5,** 23–28 (1979).
22. O'Farrell, P. H. *J. biol. Chem.* **250,** 4007–4021 (1975).
23. Bonner, W. H. & Laskey, R. A. *Eur. J. Biochem.* **46,** 83–88 (1974).
24. Bullard, B., Bell, J. Craig, R. & Leonard, K. *J. Muscle Res. Cell Motility* **1,** 194–195 (1980).
25. Lindsley, D. L. & Grell, E. H. (eds) *Genetic Variations of Drosophila melanogaster* (Carnegie Institution, Washington DC, 1968).

An adenovirus mutant defective in splicing RNA from early region 1A

David Solnick

Cold Spring Harbor Laboratory, Cold Spring Harbor,
New York 11724, USA

Many primary products of transcription in eukaryotes contain sequences that are removed during maturation of mRNAs, although the structural determinants that direct this highly specific splicing reaction are unknown. Comparative sequence studies of known splice points[1–4] have identified a short consensus sequence to which all known splice points conform to varying degrees. However, there is no direct evidence that this consensus sequence is necessary for splicing to occur. In this communication, an adenovirus mutant is described which contains a *cis* dominant defect in the splicing of RNA from one of the early transcriptional units. The defect is ascribed to changes in two adjacent nucleotides within a splice site consensus sequence. The mutations lie in a region of the adenovirus genome (the transforming region 1A) which normally produces a set of three co-terminal mRNAs differing only by the internal sequences removed during post-transcriptional processing. The mutant has thus been used not only to define sequences required for splicing, but also to investigate the pathways along which a series of related mRNAs is generated.

Isolation of the adenovirus type 5 (Ad5) mutant used a screening procedure based on differential viral growth between two human cell lines, 293 and HeLa. The 293 cell line was derived from embryonic kidney cells transformed by fragmented Ad5 DNA[5]. The cells express the left-hand 11% of the adenovirus genome[6,7] and unlike HeLa cells, support the growth of mutants mapping in this region[8,9]. The mutants fall into two complementation groups which map to the two transcriptional units in the transforming region, early regions 1A and 1B. To enhance the frequency with which additional such host range (*hr*) mutants could be isolated, a region-specific mutagenesis technique was adopted. This technique took advantage of a phenotypically wild-type variant of Ad 5 (*dl*309) which has a unique *Xba*I restriction site within early region 1A (E1A)[9]. Cleavage of *dl*309 DNA with *Xba*I yields two fragments, a small fragment containing most of E1A and a large fragment containing the remaining 96% of the genome (Fig. 1*e*). The small fragment was mutagenized with nitrous acid[10] and ligated to the untreated large fragment, and the mixture was used to transfect 293 cells[8]. Virus from the resulting plaques was screened for the host-range phenotype. The duration of nitrous acid treatment, 80–120 min, was chosen such that virus from ~20% of the plaques had the mutant phenotype.

One of the mutants, *hr*440, was characterized in detail, with particular attention to the transcription of early region 1A. E1A RNA isolated 9 h after infection of HeLa cells with Ad5, *dl*309 or *hr*440 was examined using the Berk–Sharp Method[11]. Figure 1*a* displays the three separate coding sequences protected from S_1 nuclease digestion by wild-type E1A early RNA[12]: the 610-nucleotide fragment specific for the 1.0-kilobase mRNA, the 475-nucleotide fragment specific for the 0.9-kilobase mRNA, and the 375-nucleotide fragment common to both mRNAs. Using RNA from *hr*440-infected cells, protected fragments comigrating with the 610 and 375 nucleotide bands are again present, but the 475-nucleotide band cannot be detected, even after prolonged exposure of the autoradiograph (Fig. 1*b*). The absence of the 0.9-kilobase mRNA in *hr*440-infected cells has been confirmed by fractionation of the S_1-resistant RNA–DNA hybrids on a neutral gel (not shown), and by directly displaying the RNA on a formaldehyde-containing agarose gel (Fig. 1*c*). Furthermore, the 1.0-kilobase but not the 0.9-kilobase mRNA was detected, using the Berk-Sharp method, in nuclear RNA from *hr*440 -infected cells. An additional E1A RNA species was seen in both *dl*309- and *hr*440-infected cell nuclei (data not shown). This RNA is 1.1 kilobases long, the size expected of the unspliced precursor to the E1A mRNAs[7,18]. These results suggest that the primary E1A transcript cannot be spliced to form the 0.9-kilobase mRNA, but that splicing to the 1.0-kilobase mRNA occurs normally.

A mutation could affect splicing either (1) in *trans*, through a diffusible RNA or protein responsible for the splicing event, or (2) in *cis*, through a direct effect on the DNA template or RNA precursor. To distinguish between these possibilities, S_1 analysis was performed using RNA produced by HeLa cells co-infected with *hr*440 and *dl*313, an Ad5 deletion mutant lacking most of early region 1B and a small portion (310 bases) of the 3′ end of E1A. *dl*313 complements mutants mapping in E1A but not those mapping in E1B[9,13]. Because the deletion extends partially into E1A, S_1-treated hybrids formed between *dl*313 RNA and E1A DNA migrate faster in a neutral gel than hybrids formed with *hr*440 RNA. RNA derived from each virus in the co-infected cells can therefore be distinguished, so it is possible to determine any effect the co-infection has on the production of individual viral mRNA species. If the mutation were to act in *cis*, the co-infected cells would harbour the same set of E1A mRNAs as the two single infections. If the mutation were *trans*-acting, either the *hr*440 0.9-kilobase mRNA would be present (a recessive mutation), or the *dl*313 equivalent of the 0.9-kilobase mRNA would be absent (a dominant mutation). The results of this experiment (not shown) indicate that co-infection has no effect on the production of the individual viral E1A mRNAs; thus, the *hr*440 1.0-kilobase mRNA is present and the 0.9-kilobase mRNA is absent, while the *dl*313-derived equivalents of both these mRNAs are present. The mutation acts in *cis*.

The absence of the 0.9-kilobase mRNA in *hr*440-infected cells, coupled with the *cis*-acting nature of the *hr*440 mutation, suggest the possibility that one or both of the splice points used in the synthesis of this mRNA has undergone a mutational change. The 3′ splice point, however, is common to both the 0.9 and 1.0-kilobase mRNAs, the latter of which is synthesized by the mutant. It therefore seems likely that the mutation maps near the 5′ splice site for the 0.9-kilobase mRNA. DNA sequence analysis in this region (Fig. 2*a*) uncovered two such nucleotide changes. The two mutations are adjacent, and lie

0028-0836/81/240508—03$01.00

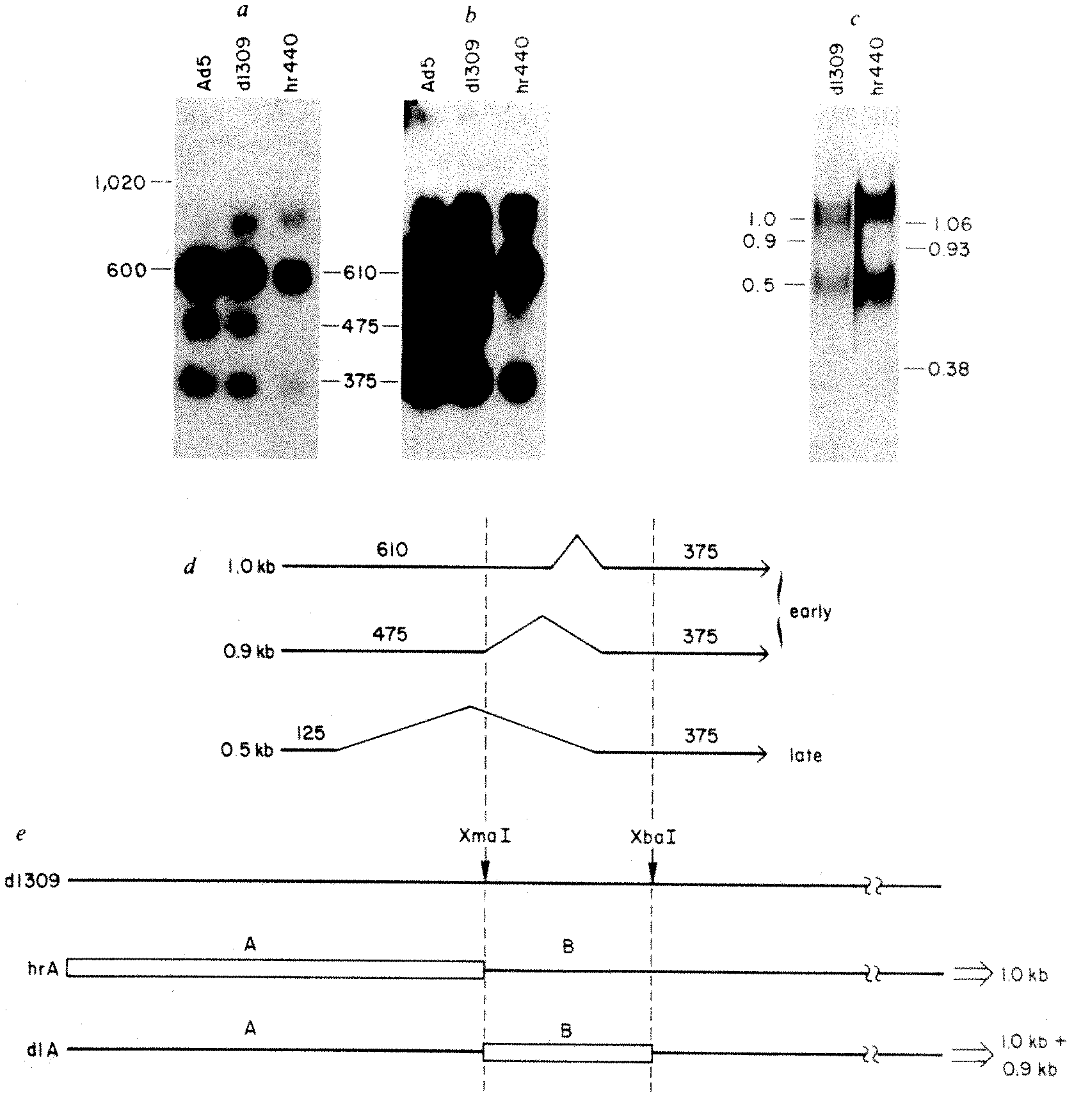


Fig. 1 Analysis of the region 1A RNAs from *hr*440 and its derivatives. *a*, Alkaline agarose gel fractionation of S_1-treated RNA–DNA hybrids. The poly(A)-containing fraction from 500μg of cytoplasmic RNA was hybridized to 0.1 μg of a plasmid (pHEB1) containing region 1A (map coordinates 0.7–4.3) inserted into pBR322. Fragments fractionated on agarose gels were transferred to nitrocellulose[23], hybridized with ^{32}P-labelled Ad5 DNA[24] and autoradiographed. Preparation of cytoplasmic RNA, RNA–DNA hybridization, S_1 treatment and gel fractionation were performed using the methods of Berk and Sharp[11]. The lengths in nucleotides of the S_1 resistant DNA fragments are shown on the right. Length markers (Ad2 *Hpa*I F and G fragments) are noted on the left. The fragment migrating more slowly than the 610-nucleotide fragment is derived from an RNA species indistinguishable from the 1.0-kilobase mRNA except for a 140-nucleotide extension at the 5′ end[25]. *b*, Longer exposure of *a*. *c*, Formaldehyde–agarose gel fractionation of poly(A)-containing cytoplasmic RNA. The poly(A)-containing fraction from 300 μg of total cytoplasmic RNA was loaded on to a 1.5% agarose gel containing 6% formaldehyde. After electrophoresis, the RNA was transferred to a nitrocellulose filter (B. Seed and D. Goldberg, personal communication) and hybridized with ^{32}P-labelled pHEB1 DNA. Lengths of the individual RNA species are marked on the left; DNA size markers are indicated on the right. The DNA markers migrate slightly faster than RNA fragments of comparable length. The lane marked *hr*440 was exposed 10 times longer than the lane marked *dl*309. *d*, Diagram of region 1A mRNAs. Caret symbols bracket the sequences removed during mRNA synthesis; arrowheads represent the 3′ ends of the mRNAs. Horizontal lines denote the sequences retained in the mRNA; lengths of these sequences are given in numbers of nucleotides. Total lengths of mRNAs are given in kilobases (kb). On the right is noted the stage after infection at which these mRNAs normally appear (from refs[12,18–20]). *e*, Structure of recombinants generated from *hr*440 and *dl*309. Open boxes denote sequences derived from *hr*440; horizontal lines denote sequences derived from *dl*309. The *Xma*I and *Xba*I sites used for the constructions are indicated in proper alignment with the RNA diagrams above. The precise location of the *Xma*I site is shown in Fig. 2*a*. On the right are noted the E1A mRNA species generated by the recombinants.

within the intron at distances of 5 and 6 nucleotides from the 5′ splice site for the 0.9-kilobase mRNA.

This result suggests that these altered nucleotides are important to the splicing process. Lerner *et al.*[3] and Rogers and Wall[4] have recently proposed a model to explain the mechanism of splicing, based on RNA–RNA hybridization between the abundant small nuclear RNA U1 and the consensus sequences at the two splice junctions of a mRNA precursor. The two mutations introduced into *hr*440 lie within the consensus sequence at the 5′ splice site for the 0.9-kilobase mRNA (Fig. 2*b*). The changes might therefore be expected to disrupt the proposed hybridization with U1 RNA. Calculation of the predicted hybrid stability[14] indicates that the negative free energy contribution of the RNA–RNA hybrid formed between U1 RNA and the wild-type 5′ splice junction is ~5.7 kcal mol^{-1}. Interestingly, this value is decreased to 3.6 kcal mol^{-1} in the mutant. Whether this difference is sufficient in itself to explain the phenotype of the mutant is unclear. The lack of information concerning the *in vivo* configuration of these putative hybrids should be emphasized.

The data presented thus far suggests that the integrity of sequences adjacent to the splice junction is required for splicing to occur. This conclusion, though, must be accompanied by one important caveat. The sequence of the entire intron for the 0.9-kilobase mRNA, along with 125 bases of 5′-flanking coding sequence, has been determined. In the coding region, two additional single-base changes were found in *hr*440, located 46 and 115 nucleotides upstream from the 5′ splice junction. Another base change was found 115 nucleotides downstream from the splice junction, roughly in the middle of the intron. To rule out the involvement of these nucleotide changes in the splicing aberration, the isolation of a phenotypically wild-type revertant of *hr*440 would be required. Repeated attempts to isolate such a revertant have been unsuccessful, presumably because more than one base change produces the mutant phenotype. To circumvent this obstacle, an alternative strategy was adopted. Viral recombinants were constructed in which a portion of the mutant or parental E1A region was replaced by a corresponding DNA segment from the parent or mutant, respectively. This is essentially a mapping procedure similar to marker rescue[15], except that the recombination is done *in vitro* rather than *in vivo*. In one recombinant (*hr*A), only the sequences including and upstream from the 5′ splice junction for the 0.9-kilobase mRNA are derived from the mutant; the remaining sequences are wild-type. In the reciprocal recombinant (*dl*A), the splice junction and upstream sequences are derived from the parental virus, while the remainder of the intron and some 3′-flanking sequences are derived from *hr*440 (Fig. 1*e*).

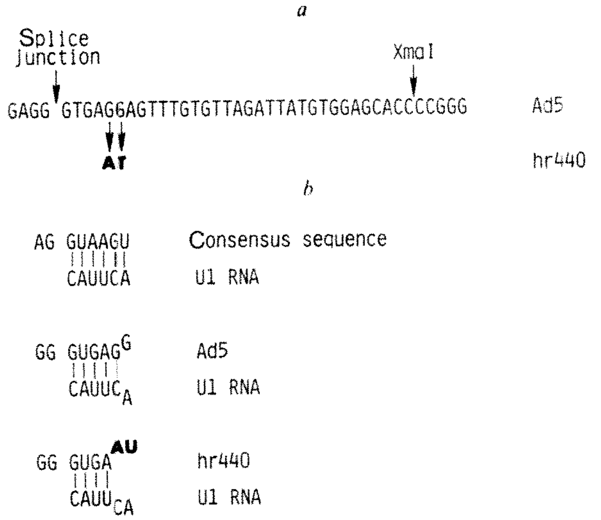


Fig. 2 DNA sequence near the 5′ splice site for the 0.9-kilobase mRNA. *a*, Ad5 sequence[27] and the nucleotide changes in *hr*440. The *hr*440 sequence was obtained by the method of Maxam and Gilbert[28], using fragments end labelled at the *Xma*I site. The position of the *Xma*I cleavage site relative to the splice junction is indicated. *b*, Putative RNA–RNA hybrids formed between U1 RNA and sequences adjacent to the 5′ splice sites. The mutations in *hr*440 are represented in boldface type. U1 RNA and consensus sequences are taken from refs 3 and 26.

The constructions were performed as follows. The small *Xba*I fragment, isolated from both *hr*440 and *dl*309 DNA, contains a single *Xma*I site located 35 bases downstream from the 5′ splice site (Fig. 2*a*). Cleavage at this restriction site produced an *Xma*I A fragment spanning from the left terminus of the genome to the *Xma*I site, and a B fragment extending from the *Xma*I site to the *Xba*I site near the right end of region 1A (Fig. 1*e*). The *Xma*I cleavage products were separated and used in two ligation reactions, each containing the large *Xba*I fragment from *dl*309 and either (1) the *hr*440 *Xma*I A fragment and the *dl*309 *Xma*I B fragment, or (2) the *dl*309 *Xma*I A fragment and the *hr*440 *Xma*I B fragment. The ligation mixtures were used to transfect 293 cells, and virus from the resultant plaques was analysed for host range phenotype and synthesis of E1A RNAs.

Virus derived from ligation 1 (*hrA*) gave a plaquing ratio (titre on 293/titre on HeLa) of 10^4, which is indistinguishable from that of *hr*440. Virus derived from ligation 2 (*dl*A) showed a plaquing ratio of 70, compared with a ratio of 1 for the parental virus *dl*309. The residual host range exhibited by *dl*A may be due to the mutation located downstream from the 5′ splice site which was detected by sequence analysis (this mutation converts an arginine codon to a lysine codon in the coding sequence of the 1.0-kilobase mRNA), or to other unidentified mutation(s) further downstream. Importantly, *hr*A, like *hr*440, produced the 1.0-kilobase mRNA but no detectable levels of the 0.9-kilobase mRNA on infection of HeLa cells; *dl*A, on the other hand, produced both the 1.0- and 0.9-kilobase mRNAs (data not shown). The mutation(s) responsible for the splicing defect thus maps to the left of the *Xma*I site. The only mutational changes in this region which lie in or near the intron for the 0.9-kilobase mRNA are those shown in Fig. 2*a*. The possibility that the upstream base changes contribute to the splicing defect has not been eliminated but is rendered unlikely by two studies on the splicing of RNA produced from templates deleted of sequences upstream from a 5′ splice site. One study showed that the presence of only 18 bases upstream from an intron in the β^{maj} globin gene is sufficient for proper removal of that intron by splicing[16]. In the second study, synthesis of an SV40 early mRNA occurred in cells infected by a deletion mutant lacking the region 29–104 nucleotides upstream from the 5′ splice junction[17].

At late times after infection, adenovirus produces still another mRNA from region 1A. The structure of this 0.5-kilobase mRNA, as determined by electron microscopy[18], is shown in Fig. 1*d*. Comparison of the three E1A mRNAs indicates that they constitute a set of co-terminal RNAs differing only by the positions of the 5′ splice sites. Analogous configurations exist in early region 1B mRNAs[12,18–20], and in the SV40 early mRNAs[21,22]. Two simple pathways for the synthesis of these mRNAs can be envisioned: (1) a parallel pathway, in which all three mRNAs are derived directly from an unspliced precursor, and (2) a sequential pathway, where each mRNA is derived as a product of stepwise removal of intervening sequences in the order 1.0 → 0.9 → 0.5 kilobase mRNA. More complex combinations of these two pathways are also possible. The two models make mutually distinct predictions regarding the synthesis of E1A mRNAs, some of which can be tested with *hr*440. If the parallel pathway is used, the 0.5-kilobase mRNA will be present in *hr*440-infected cells even though the 0.9-kilobase mRNA is absent. If the pathway is sequential, the 0.5-kilobase mRNA will be absent.

The presence or absence of the 0.5-kilobase mRNA was investigated using an RNA gel blotting procedure. This procedure was required because, for unknown reasons, the Berk–Sharp method does not detect the 0.5-kilobase mRNA. Cytoplasmic RNA isolated 17 h post-infection was fractionated on a formaldehyde-containing agarose gel and transferred to a nitrocellulose filter. E1A-specific RNA was illuminated by hybridization to a ^{32}P-labelled E1A probe. The result of this experiment (Fig. 1*c*) clearly indicates the presence of the 0.5-kilobase mRNA in *hr*440-infected cells. The splicing pathway in therefore not obligatorily sequential. However, the possibility remains that the 1.0- and 0.9-kilobase mRNAs are normally synthesized sequentially, whereas the 0.5-kilobase mRNA is derived in parallel from either the 1.0-kilobase mRNA or the unspliced precursor. Biogenesis of the SV40 early mRNAs also does not require a sequential pathway. Synthesis of the smaller of the two co-terminal early mRNAs (the 16S mRNA) is not affected in cells infected with deletion mutants unable to synthesize the larger (19S) mRNA[17,29,30].

Thus we conclude that sequences located within the intron of region 1A and very near the 5′ splice junction are required for RNA processing at this splice junction. Furthermore, biogenesis of the three overlapping E1A mRNAs does not proceed in a simple sequential manner. Because the smallest of the three mRNAs is synthesized in the absence of the next larger, this larger RNA cannot be a necessary precursor to the smaller.

I thank Joe Sambrook for advice and critical reading of the manuscript. This work was funded by NCI grant CA 13106.

Received 7 January; accepted 16 March 1981.

1. Breathnach, R., Benoist, C., O'Hare, K., Gannon, F. & Chambon, P. *Proc. natn. Acad. Sci. U.S.A.* **75,** 4853–4857 (1978).
2. Seif, I., Khoury, G. & Dhar, R. *Nucleic Acids Res.* **6,** 3387–3398 (1979).
3. Lerner, M. R., Boyle, J. A., Mount, S. M., Wolin, S. L. & Steitz, J. A. *Nature* **283,** 220–224 (1980).
4. Rogers, J. & Wall, R. *Proc. natn. Acad. Sci. U.S.A.* **77,** 1877–1879 (1980).
5. Graham, F. L., Smiley, J., Russell, W. C. & Nairn, R. *J. gen. Virol.* **36,** 59–72 (1977).
6. Aiello, L., Huebner, K. & Weinmann, R. *Virology* **94,** 460–469 (1979).
7. Berk, A. J., Lee, F., Harrison, T., Williams, J. & Sharp, P. A. *Cell* **17,** 935–944 (1979).
8. Frost, E. & Williams, J. *Virology* **91,** 39–50 (1978).
9. Jones, N. & Shenk, T. *Cell* **17,** 683–689 (1979).
10. Williams, J. F., Gharpure, M., Ustacelebi, S. & McDonald, S. *J. gen. Virol.* **11,** 95–101 (1971).
11. Berk, A. J. & Sharp, P. A. *Cell* **12,** 721–732 (1977).
12. Berk, A. J. & Sharp, P. A. *Cell* **14,** 695–711 (1978).
13. Ross, S. R., Levine, A. J., Galos, R. S., Williams, J. & Shenk, T. *Virology* **103,** 475–492 (1980).
14. Salser, W. *Cold Spring Harb. Symp. quant. Biol.* **42,** 985–1002 (1977).
15. Lai, C.-J. & Nathans, D. *Virology* **60,** 466–475 (1974).
16. Hamer, D. A. & Leder, P. *Cell* **17,** 737–747 (1979).
17. Khoury, G., Gruss, P., Dhar, R. & Lai, C.-J. *Cell* **18,** 85–92 (1979).
18. Chow, L. T., Broker, T. R. & Lewis, J. B. *J. molec. Biol.* **134,** 265–303 (1979).
19. Kitchingman, G. R. & Westphal, H. *J. molec. Biol.* **137,** 23–48 (1980).
20. Perricaudet, M., Le Moullec, J. M. & Pettersson, U. *Proc. natn. Acad. Sci. U.S.A.* **77,** 3778–3782 (1980).
21. Berk, A. J. & Sharp, P. A. *Proc. natn. Acad. Sci. U.S.A.* **75,** 1274–1278 (1978).
22. Reddy, V. B., Ghosh, P. K., Lebowitz, P., Piatak, M. & Weissman, S. M. *J. Virol.* **30,** 279–296 (1979).
23. Southern, E. M. *J. molec. Biol.* **98,** 503–517 (1975).
24. Maniatis, T., Kee, S. G., Efstradiatis, A. & Kafatos, F. C. *Cell* **8,** 163–182 (1976).
25. Sambrook, J. *et al. Cold Spring Harb. Symp. quant. Biol.* **44,** 569–584 (1980).
26. Reddy, R., Ro-Choi, T. S., Henning, D. & Busch, H. *J. biol. Chem.* **249,** 6486–6494 (1974).
27. van Ormondt, H., Maat, J., De Waard, A. & Van der Eb, A. J. *Gene* **4,** 309–328 (1978).
28. Maxam, A. & Gilbert, W. in *Meth. Enzym.* **65,** 499–560 (1980).
29. Thimmappaya, B. & Shenk, T. *J. Virol.* **30,** 668–673 (1979).
30. Volckaert, G., Feunteun, J., Crawford, L., Berg, P. & Fiers, W. *J. Virol.* **30,** 674–682 (1979).

In vitro synthesis of polypeptides encoded by the long terminal repeat region of mouse mammary tumour virus DNA

Clive Dickson, Rosalind Smith & Gordon Peters

Imperial Cancer Research Fund Laboratories, PO Box 123,
Lincoln's Inn Fields, London WC2A 3PX, UK

The integrated proviral DNA of RNA tumour viruses is bounded by long terminal repeat (LTR) segments of several hundred base pairs which result from duplication of sequences from both the 5′ and 3′ ends of the viral genome RNA[1-3]. Recently, DNA sequence analysis of the LTRs of several retroviruses has indicated that this region of the provirus may contain regulatory functions for viral RNA transcription [4-8]. In this connection, mouse mammary tumour virus (MuMTV) is of particular interest in that transcription of the viral RNA can be specifically modulated by glucocorticoid hormones, suggesting that the site of steroid action may be contained within the LTR[9]. Moreover, the LTR of MuMTV is unusually long, extending for about 1,300 base pairs. We now present data which suggest that, in contrast with other retroviruses, the MuMTV LTR includes approximately 1,000 nucleotides in an open reading frame, capable of encoding a series of polypeptides which are quite distinct from any of the known structural components of MuMTV virions. These proteins can be expressed *in vitro* from cloned, recombinant DNA in which the MuMTV LTR has been inserted into the ampicillin-resistance gene of a bacterial plasmid. Tryptic peptide mapping has confirmed that these products are identical to the analogous series of proteins identified by *in vitro* translation of polyadenylated fragments of MuMTV genome RNA[10].

The ability of certain retroviruses to transform cells in culture and induce rapid neoplasia *in vivo* can be attributed to the expression of specific proteins encoded by sequences within the viral genome. In many instances, these sequences have been shown to be closely related to and possibly derived from analogous elements in normal uninfected cells[11,12]. This concept cannot, however, fully account for neoplasia induced by the large number of retroviruses, such as MuMTV, which do not seem to transform cells *in vitro* and for which no transformation-specific protein or sequences have yet been identified. To approach this question, we have recently examined the total protein-coding capacity of the genomic RNA by *in vitro* translation and found, as well as the expected translation products derived from the viral structural genes, *gag*, *pol* and *env*, a series of overlapping proteins whose tryptic peptide maps were quite distinct from those of the known structural components of the virus[10]. The largest of these (estimated molecular weight MW 36,000 (36K)) would require a coding sequence of about 1,000 nucleotides in a continuous open reading frame. Peptide mapping and limited sequence analysis of the amino-terminal peptides indicated that the smaller components of the series were probably initiated at different methionine codons within the same open reading frame. As no function has been ascribed to these non-structural proteins, we propose to use the term *orf* as an abbreviation for the open reading frame sequences and the products they encode. However, this is not meant to imply that *orf* represents an additional gene.

Translation of subgenomic fragments of MuMTV RNA indicated that the *orf* sequences must be located within the 3′-terminal 2,000 nucleotides of the genome RNA[10]. To map the open reading frame more precisely, we have examined the protein-coding capacity of selected regions of MuMTV proviral DNA cloned into bacterial plasmid vectors. As with other retroviruses, the genome of MuMTV contains regions at both 5′ and 3′ termini (designated U5 and U3, respectively) which are duplicated during formation of the integrated proviral DNA[3] to give a provirus whose structure resembles that of a bacterial transposable element in that the coding sequences for the *gag*, *pol* and *env* genes are flanked at each end by identical terminal repeats containing both U3 and U5 (Fig. 1)[4-7,13]. In MuMTV, these LTRs are about 1,300 nucleotides long, of which ~1,200 correspond to the 3′ end of the genome RNA (U3)[13,14]. Cleavage of MuMTV proviral DNA with the restriction endonuclease *Pst*I generates a 1.4-kilobase DNA fragment containing all but eight nucleotides of the left-hand LTR. Using derivatives of a recombinant bacterial plasmid containing this DNA fragment inserted at the *Pst*I site of the drug resistance vector pBR322, we have now examined the protein-coding capacity of the LTR DNA in an *in vitro* system.

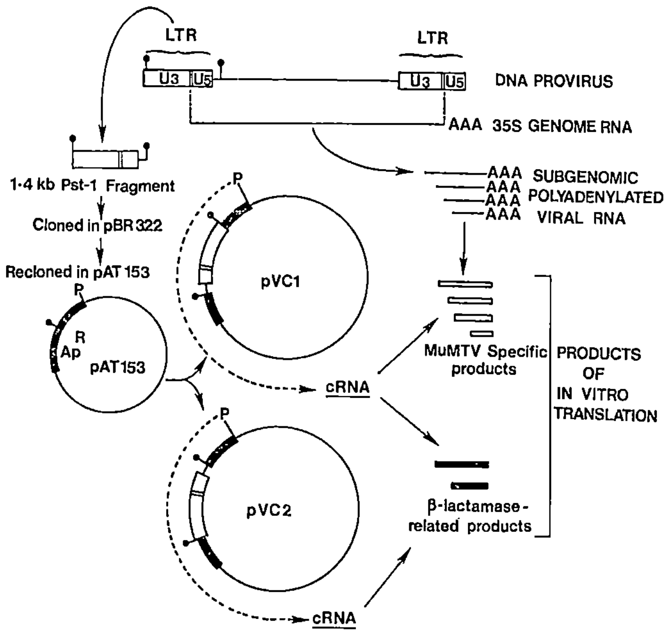


Fig. 1 Construction and *in vitro* expression of recombinant plasmids containing MuMTV DNA. A recombinant plasmid containing the 1.4-kilobase *Pst*I fragment derived from MuMTV proviral DNA inserted into the *Pst*I site of the drug resistance vector pBR322 was originally constructed by Majors and Varmus[13]. The 1.4-kilobase fragment was excised from the plasmid, purified by agarose gel electrophoresis and ligated into the *Pst*I site of the non-mobilizable plasmid pAT153 following standard procedures. Recombinant plasmids were propagated in *E. coli* HB101 and screened for the presence of a single 1.4-kilobase insert of MuMTV DNA. The orientation of the fragment with respect to the promoter (P) for the ampicillin resistance gene (Ap$^{\mathrm{R}}$) was determined by digestion with appropriate restriction enzymes and two representative examples, pVC1 and pVC2, were chosen for further study. Purified plasmid DNAs were used to direct *in vitro* transcription by purified *E. coli* RNA polymerase and the protein-coding potential of the resultant cRNA was assessed by translation in a rabbit reticulocyte lysate as described in Fig. 2 legend.

In pBR322, and its non-mobilizable derivative pAT153, the cleavage site for the enzyme *Pst*I lies in the coding sequences for the β-lactamase enzyme which confers resistance to ampicillin on recipient bacteria[15,16]. Transcription of the β-lactamase gene is thought to initiate at a promoter site ~545 nucleotides upstream from the *Pst*I site[15,17]. We considered that RNA transcripts initiated at this promoter could be extended beyond the *Pst*I site to include any foreign sequences inserted at this position provided that they do not contain transcriptional termination signals. Translation of the resultant RNA in a cell-free system may therefore yield products derived from both the residual β-lactamase gene and any segment of open reading frame within the inserted DNA. Knowing the nucleotide sequence of pAT153, we calculate that the residual part of the β-lactamase gene, from the initiation codon to the *Pst*I site,

0028-0836/81/240511—03$01.00

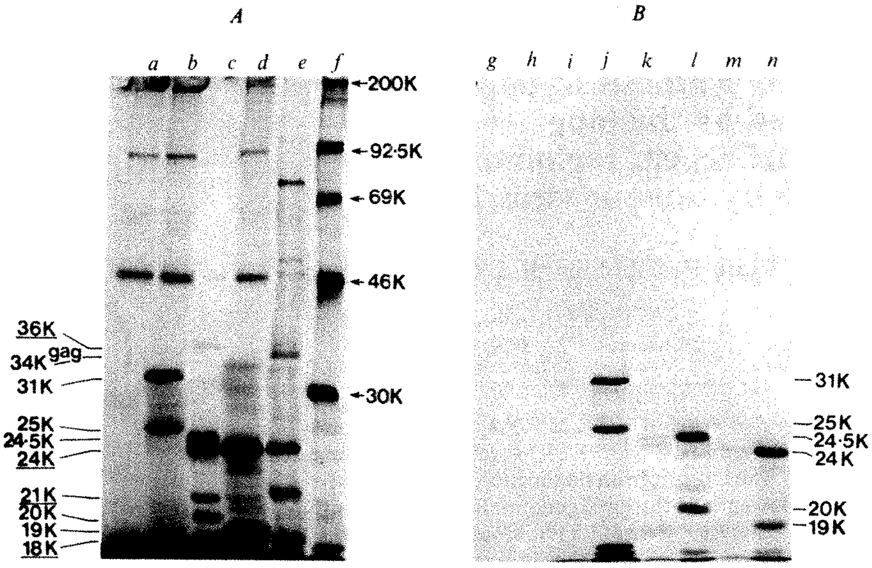


Fig. 2 Analysis of products derived by *in vitro* expression of recombinant plasmid DNAs. *A*, Purified plasmid DNAs were used as templates for *in vitro* RNA synthesis in reactions containing 20 mM HEPES (*p*H 7.9), 100 mM NaCl, 10 mM Mg acetate, 1 mM $MnCl_2$, 5 mM dithiothreitol, 2.6 mM each of ATP, CTP, GTP and UTP, 1.2 µCi ^{3}H-UTP (Radiochemical Centre, 52 Ci mmol^{-1}), 100 µg ml^{-1} plasmid DNA and 10–20 units of *E. coli* RNA polymerase (New England Biolabs) in a total volume of 0.12 ml (ref. 19). Following incubation at 37 °C for 1 h, the products of the reaction were phenol extracted and ethanol precipitated, and portions taken for *in vitro* translation. Cell-free protein synthesis was carried out in nuclease-treated rabbit reticulocyte lysates primed with 1–2 µg of plasmid-derived cRNA or MuMTV genome RNA as previously described[10,18]. The ^{35}S-methionine-labelled products were analysed by electrophoresis in 9% polyacrylamide gels, in the presence of SDS, and visualized by autoradiography: *a*, no RNA; *b*, pAT153 cRNA; *c*, pVC1 cRNA; *d*, pVC2 cRNA; *e*, MuMTV viral RNA; *f*, MW standards. The approximate MWs ($\times 10^3$) of the major products are indicated as computed from their electrophoretic mobilities relative to the standards. Numbers underlined indicate those products not-related to β-lactamase *B*, Aliquots of translation reactions *a*–*d* were immune precipitated using either normal rabbit serum shown in tracks *g*, *i*, *k* and *m*, or anti-β-lactamase serum (anti-TEM-1) as in tracks *h*, *j*, *l* and *n*. Immune complexes were recovered by absorption to *Staphylococcus aureus*, dissociated and the labelled proteins analysed as described above. *g*, *h*, No RNA; *i*, *j*, pAT153 cRNA; *k*, *l*, pVC1 cRNA; *m*, *n*, pVC2 cRNA.

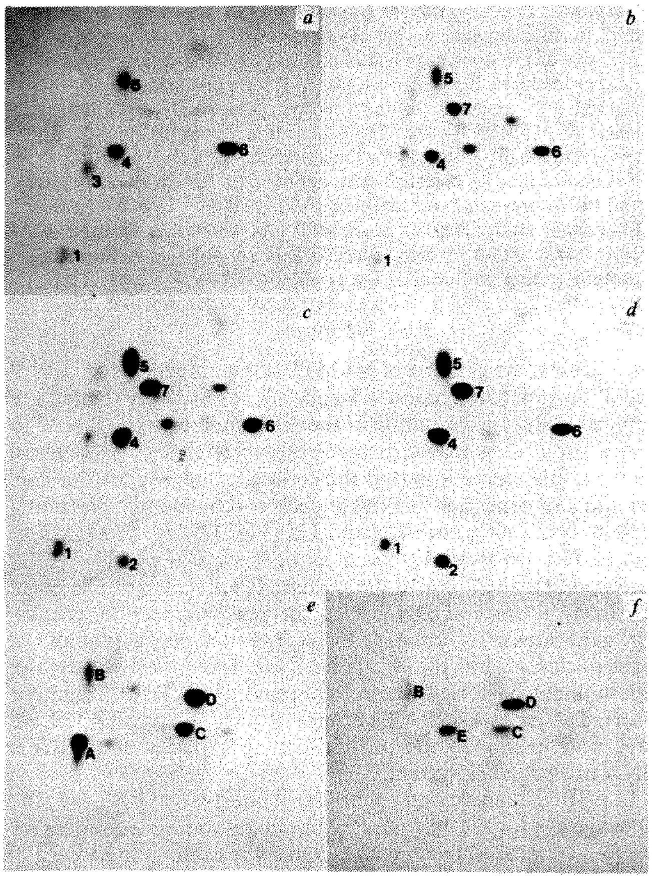


Fig. 3 Tryptic peptide mapping of methionine-labelled products translated from plasmid DNA and MuMTV viral RNA. Polypeptide products identified as bands on the autoradiograph shown in Fig. 2*A* were excised, eluted from the gel and subjected to performic acid oxidation and trypsin digestion as previously described[10]. The labelled peptides were then separated in two dimensions by electrophoresis and chromatography, and visualized by autoradiography[10]. *a*, 36K product from pVC1 cRNA; *b*, 24K product from pVC1 cRNA; *c*, mixture of equivalent amounts of the 24K products shown in *b* and *d*; *d*, 24K product from MuMTV viral RNA; *e*, β-lactamase-related 24.5K product from pVC1 cRNA; *f*, β-lactamase-related 31K product from pAT153 cRNA.

could encode a protein of MW at least 23.9K as compared with 31K for the authentic, full-sized product[15]. The precise size of such a truncated product would depend on the location of the first in-frame termination signal encountered in the inserted DNA. Using MuMTV DNA sequences originally cloned in pBR322, two pAT153-based recombinants have been constructed in which the LTR is in a 5′ to 3′ orientation (pVC1) or vice versa (pVC2) with respect to the direction of transcription from the β-lactamase promoter[15]. Purified plasmid DNA was used to direct *in vitro* RNA synthesis using *Escherichia coli* RNA polymerase. Initiation at alternative promoters in the plasmid DNA was suppressed by including 0.1 M KCl in the transcription reactions[17]. In the conditions used, the yield of complementary RNA (cRNA) per µg of input plasmid DNA was 3:1, suggesting several rounds of transcription. The cRNA produced was then purified by phenol extraction and used to programme *in vitro* protein synthesis in nuclease-treated rabbit reticulocyte lysates as previously described[10,18]. The ^{35}S-methionine-labelled products were analysed by SDS-acrylamide gel electrophoresis and tryptic peptide mapping following standard procedures[10].

The coding capacities of pVC1 and pVC2 DNAs were compared with those of the parental plasmid pAT153 and MuMTV viral RNA (Fig. 2). pAT153 DNA, containing no insert at the *Pst*I site, directed the synthesis of two major polypeptides (Fig. 2*A*). The larger of these, with an estimated molecular weight of 31K, was assumed to be the authentic primary translation product of the β-lactamase gene, while the smaller 25K product probably reflected initiation of translation at an internal methionine codon within the β-lactamase gene. Both these *in vitro* products could be immune precipitated with an antiserum against the purified protein (anti-TEM-1 serum) (Fig. 2*B*), confirming that they are indeed related to β-lactamase. In contrast to the parental plasmid, pVC1 DNA directed the synthesis of six major polypeptides of which only two (MWs 24.5K and 20K) were related to β-lactamase as determined by immune precipitation. Thus, the other products, designated 36K, 24K, 21K and 18K, were probably encoded by the MuMTV DNA sequences in the plasmid. As they could not be precipitated by the anti-β-lactamase serum, they probably represent the products of independent, internal initiation events within additional open reading frame sequences provided by the insertion of MuMTV DNA. In pVC1, the inserted MuMTV DNA is in the correct orientation for transcription of the positive sense strand from the β-lactamase promoter. With pVC2 DNA, where the LTR is in the reverse orientation, no MuMTV specific products were detected and the two major polypeptides obtained (MWs 24K and 19K) were both immune precipitated by anti-β-lactamase serum. Thus, the negative sense strand of the MuMTV DNA contains no significant stretches of open reading frame detectable in the *in vitro* expression system.

An important feature of the additional products encoded by pVC1 DNA was that their size correlated almost exactly with that of members of the *orf* series of proteins translated from MuMTV viral RNA (Fig. 2*Ae*). The viral RNA also programmed the synthesis of products from the structural genes, including a 34K *gag* peptide which migrates very close to the 36K *orf* product[10]. However, as no antiserum against the *orf* proteins is available, we could not verify this relationship by immune precipitation. To circumvent this problem, the various products obtained by *in vitro* expression of plasmid DNA were eluted and further analysed by tryptic peptide mapping. As expected, the fingerprints of the 36K, 24K, 21K and 18K products from pVC1 represented an overlapping series in that most peptides were shared by each member of the series (see Fig. 3*a*, *b*). Peptide analysis also indicated that the pVC1-encoded proteins were virtually identical to those of the corresponding *orf* products obtained by translation of viral RNA (Fig. 3*c*, *d*)[10], and substantiated the previous indications that the immune-precipitable products observed in Fig. 2*B* were derived from the β-lactamase gene (Fig. 3*e*, *f*).

We therefore conclude that approximately 1,000 of the 1,300 nucleotides of the MuMTV provirus LTR are organized into a continuous open reading frame which we have termed *orf*. The exact location of the *orf* sequences in relation to the ends of the cloned fragment of MuMTV DNA used in these studies is unknown, but several preliminary conclusions can be drawn. As the *orf* proteins were originally detected by *in vitro* translation of the 3′ end of the viral RNA, the coding sequences cannot include the U5 portion of the LTR nor the additional sequences distal to U5 in the cloned *Pst*I fragment, and must therefore be wholly or partly contained within the U3 segment of the genome. As the *Pst*I fragment used to construct the recombinant plasmids lacks eight nucleotides from the 5′ end of the LTR, it is possible that the *orf* sequences extend beyond the *Pst*I site and that the linkage of MuMTV to plasmid DNA has occurred within the open reading frame. Thus, initiation of the largest *in vitro* product from pVC1 cRNA may have occurred within the β-lactamase gene, but in a different reading frame, and fortuitously given rise to a 36K species. However, no major differences were observed between the methionine-labelled tryptic peptides of the LTR DNA- and viral RNA-encoded 36K products, and both contained a unique peptide not present in the 24K products or other members of the *orf* series (see Fig. 3*a*, *b* and ref. 10). This would imply that at least the major part and probably all of the 36K coding sequences are located within U3.

The U3 regions of several retroviruses, including MuMTV, have now been completely sequenced and shown to contain elements presumably involved in initiation of transcription and polyadenylation of the viral RNA[4–8,13,14]. In contrast to the other retroviruses, the MuMTV LTR was also found to include a substantial region of open reading frame. However, on the basis of the published sequence data, this reading frame would be insufficient to encode the complete 36K *orf* product, although it could encompass the 24K and smaller *orf* products[14].

Drug resistance plasmid pBR322 containing the MuMTV LTR fragment was provided by Drs J. Majors and H. E. Varmus and anti-TEM-1 serum by Dr M. Boulton (Glaxo).

Received 15 January; accepted 18 March 1981.

1. Hsu, T. W., Sabran, J. L., Mark, G. E., Guntaka, R. V. & Taylor, J. M. *J. Virol.* **28,** 810–818 (1978).
2. Shank, P. R. *et al. Cell* **15,** 1383–1395 (1978).
3. Cohen, J. C., Majors, J. E. & Varmus, H. E. *J. Virol.* **32,** 483–496 (1979).
4. Shimotohno, K., Mizutani, S. & Temin, H. M. *Nature* **285,** 550–554 (1980).
5. Van Beveren, C., Goddard, J. G., Berns, A. & Verma, I. M. *Proc. natn. Acad. Sci. U.S.A.* **77,** 3307–3311 (1980).
6. Shoemaker, C., Goff, S., Gilboa, E., Paskind, M., Mitra, S. W. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **77,** 3932–3936 (1980).
7. Dhar, R., McClements, W. L., Enquist, L. W. & Van de Woude, G. F. *Proc. natn. Acad. Sci. U.S.A.* **77,** 3937–3941 (1980).
8. Benz, E. W. Jr., Wydro, R. M., Nadal-Ginard, B. & Dina, D. *Nature* **288,** 665–669 (1980).
9. Ringold, G. M. *Biochem. biophys. Acta* **560,** 487–508 (1979).
10. Dickson, C. & Peters, G. *J. Virol.* **37,** 36–47 (1981).
11. Sharp, P. A. *Cold Spring Harb. Symp. quant. Biol.* **44,** 1305–1322 (1979).
12. Hayman, M. J. *J. gen. Virol.* **52** (in the press).
13. Varmus, H. E. & Majors, J. E. *9th a. UCLA-ICN Symp. Animal Virus Genetics* (Academic, New York, in the press).
14. Donehower, L., Huang, A. & Hager, G. L. *J. Virol.* **37,** 226–238 (1981).
15. Sutcliffe, J. G. *Proc. natn. Acad. Sci. U.S.A.* **75,** 3737–3741 (1978).
16. Twigg, A. J. & Sherratt, D. *Nature* **283,** 216–218 (1980).
17. West, R. W. Jr. & Rodriguez, R. L. *Gene* **9,** 175–193 (1980).
18. Pelham, H. R. B. & Jackson, R. J. *Eur. J. Biochem.* **67,** 247–256 (1976).
19. Sambrook, J., Sharp, P. A. & Keller, W. *J. molec. Biol.* **70,** 57–71 (1972).

Species variation in the specificity of ribulose biphosphate carboxylase/oxygenase

Douglas B. Jordan* & William L. Ogren†

* Department of Agronomy, University of Illinois, Urbana, Illinois 61801, USA
† US Department of Agriculture, Science and Education Administration-Agricultural Research, Urbana, Illinois 61801, USA

The balance between photosynthesis and photorespiration in many species, including most crop plants, is determined by the kinetic properties of ribulose-1,5-bisphosphate (RuBP) carboxylase/oxygenase[1,2]. Photosynthesis is initiated by the carboxylase activity[3] while the oxygenase activity catalyses the first reaction in the photorespiratory pathway[1,2,4]. In these reactions, CO_2 and O_2 are competitive substrates[2]. Because O_2 inhibits carboxylation, and photorespiration oxidizes reduced carbon to CO_2 with no known benefit to the plant, it has been suggested that photosynthetic efficiency and thus productivity might be increased by chemical or genetic alterations of the enzyme which increase carboxylation or decrease oxygenation[5–7]. On the other hand, it has been argued that RuBP carboxylase/oxygenase cannot completely discriminate between CO_2 and O_2, so that photorespiration is unavoidable[8]. From analyses of RuBP carboxylase/oxygenase enzymes purified from several different species, we report here large differences in specificity towards the substrates CO_2 and O_2. Evolutionary pressures seem to have directed the enzyme towards more efficient utilization of CO_2.

The ratio of carboxylation to oxygenation catalysed by RuBP carboxylase/oxygenase is described by the equation[2]:

$$\nu_c/\nu_o = (V_cK_o/V_oK_c)([CO_2]/[O_2])$$

where ν_c and ν_o are the velocities of carboxylation and oxygenation, respectively, V_c and V_o the maximal velocities of the two reactions, and K_c and K_o the Michaelis constants for CO_2 and O_2. The substrate specificity factor, V_cK_o/V_oK_c, determines the relative rates of the two reactions at any given CO_2 and O_2 concentrations; a high specificity value indicates a greater specificity for CO_2. We recently devised a method in which both enzyme activities can be assayed simultaneously[9], and which, when used at several $[CO_2]/[O_2]$ values, permits the direct determination of V_cK_o/V_oK_c. Data obtained from applying this assay to enzymes purified from six species are shown in Fig. 1. The specificity factors and relevant kinetic constants for these and enzymes from nine additional sources are summarized in Table 1. The specificity factors ranged from about 12, in enzymes isolated from anaerobic photosynthetic bacteria, to about 80, in enzymes isolated from higher plants. In other plant groups the specificity factors increased from about 48 in two cyanobacterium species to 54 in *Euglena gracilis* to 62 in two species of green algae.

0028–0836/81/240513—03$01.00

Table 1 Specificity factor and component kinetic constants for RuBP carboxylase/oxygenase purified from several species

Species	Specificity factor (V_cK_o/V_oK_c)	K_c (μM)	K_o (μM)	V_c/V_o (ratio)
C_3 plants				
Glycine max	82 ± 5	9	430	1.7
Tetragonium expansa	81 ± 1	13	600	1.8
Spinacea oleracea	80 ± 1	14	480	2.3
Lolium perenne	80 ± 1	16	500	2.6
Nicotiana tabacum	77 ± 1	11	650	1.3
C_4 plants				
Amaranthus hybridus	82 ± 4	16	640	2.0
Zea mays	78 ± 3	34	810	3.3
Green algae				
Scenedesmus obliquus	63 ± 2	38	660	3.6
Chlamydomonas reinhardii	61 ± 5	29	480	3.7
Euglena gracilis	54 ± 2	25	410	3.3
Cyanobacteria				
Aphanizomenon flos-aquae	48 ± 2	105	990	5.1
Cocochloris peniocystis	47 ± 2	121	1,220	4.7
Photosynthetic bacteria				
Rhodospirillum rubrum	15 ± 1	89	406	3.3
Rhodopseudomonas sphaeroides II	9 ± 1	80	220	3.3
Rhodopseudomonas sphaeroides I	62 ± 4	36	840	2.7

Enzymes were purified and specificity factors were determined as described in Fig. 1 legend. For K_c and K_o determinations, the proteins were preincubated at 4 °C for 4 h in 10 mM $NaH^{14}CO_3$ (2 Ci mol^{-1}), 10 mM $MgCl_2$ and 50 mM Bicine at *p*H 8.5. Assays (30 s) were initiated by adding 100 μg or less of activated protein in 20 μl to reaction mixtures at 25 °C containing 0.40 mM RuBP, 10 mM $MgCl_2$ and 50 mM Bicine at *p*H 8.22 or 8.30 in a total volume of 1.0 ml. $NaH^{14}CO_3$ concentrations varied between 0.7 and 15 mM, and O_2 concentrations were 0, 0.62 and 1.24 mM. The reactions were stopped by adding 0.5 ml of 3M formic acid in methanol. After drying the samples, $^{14}CO_2$ incorporation was determined by scintillation spectroscopy. K_c was estimated from a Scatchard plot; K_o, determined as the $K_i(O_2)$, was estimated from Dixon plots and from secondary plots of $K_i(O_2)$ (apparent) as a function of CO_2 concentration. Lines were fitted by least-squares analysis. V_c/V_o was calculated from the relationship: V_c/V_o = specificity factor × (K_c/K_o). All experiments included *S. oleracea* enzyme as a control. Standard errors in the specificity factor determinations are listed. The mean standard error for K_c was ±7% of the value, and ±12% for K_o.

Rhodopseudomonas sphaeroides contains two RuBP carboxylase/oxygenase enzymes[10]: one is a lower-molecular-weight protein which, like the enzyme from *Rhodospirillum rubrum*, consists of a single subunit type; the other is a higher-molecular-weight protein similar to the enzyme comprising eight large and eight small subunits obtained from green algae and higher plants. The low-molecular-weight protein (designated II in Fig. 1 and Table 1) was kinetically similar to the *R. rubrum* enzyme, whereas the high-molecular-weight protein (I) was kinetically similar to the algal enzymes. These correlations suggest that high CO_2 specificity requires both subunit types.

The major cause of variation in the specificity factor was changes in K_c, as reasoned previously[6]. However, substantial differences in K_o and V_c/V_o were also found (Table 1). The low-molecular-weight enzymes from photosynthetic bacteria had a low affinity for CO_2 and a high affinity for O_2. The increased specificity factor in cyanobacteria was caused by a decreased affinity for O_2. The affinity for CO_2 was greater in green algal and higher plant enzymes than in bacterial enzymes by up to an order of magnitude, while the affinity for O_2 was intermediate to that found in enzymes from photosynthetic and cyanobacteria. Variation in the ratio of the maximal velocities of the two reactions was also found, ranging from V_c/V_o values of about 2 in C_3 plants (those which fix CO_2 by the Calvin cycle reactions) to about 5 in the cyanobacteria.

These observations suggest that RuBP carboxylase/oxygenase has evolved to compensate, at least in part, for the geological shift from an atmosphere containing high CO_2 and low O_2 concentrations to one consisting of low CO_2 and high O_2 levels, a shift occasioned by the appearance of O_2-evolving photosynthesis[11]. The specificity for CO_2 in algal and cyanobacterial enzymes is not as high as that observed in enzymes from C_3 plants, perhaps because the unicellular species have developed mechanisms unknown in detail which concentrate CO_2 internally by a factor of 10 or more[12]. The resulting high internal CO_2 concentration reduces the need for a more efficient enzyme compared with C_3 species, in which the carboxylase is in relatively free equilibrium with the external atmosphere. C_4 plants (those which initially fix CO_2 by the enzyme phosphoenolpyruvate carboxylase) also concentrate CO_2 (ref. 13), yet enzymes from these plants have a high specificity factor (Table 1). C_4 species were derived from C_3 species[14], so that high CO_2 specificity of the C_3 enzyme seems to have been conserved despite the generally higher K_c values in C_4 compared with C_3 enzymes[15].

The data presented here demonstrate that the balance between RuBP carboxylation and oxygenation is not immutable, and that modifications increasing carboxylation efficiency and the affinity for CO_2 have occurred during the course of natural history. In the context of plant productivity, the question

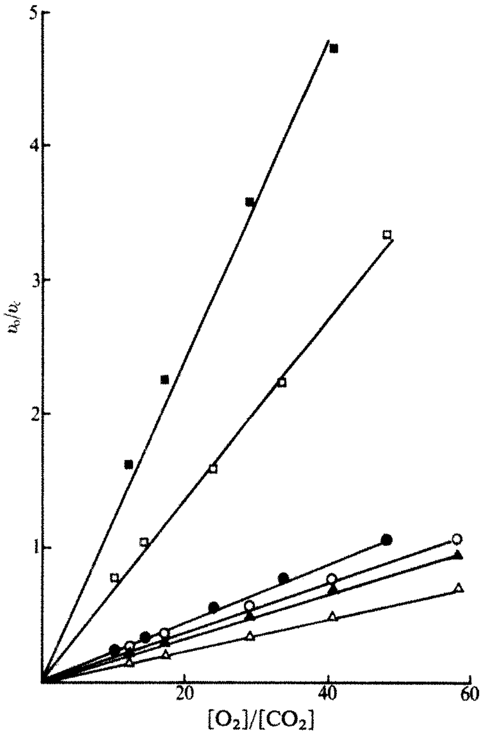


Fig. 1 Determination of the substrate specificity factor. RuBP carboxylase/oxygenase from the various species was purified by a combination of gel filtration and ion-exchange chromatography using Sepharose 6B, Sephadex G-200 and DEAE-Sephacel. Following overnight dialysis at 4 °C against 20 mM Bicine at *p*H 8.0 and 0.1 mM EDTA, RuBP carboxylase and oxygenase activities were determined simultaneously as previously described[9]. Reaction mixtures containing 2.5 mM $NaH^{14}CO_3$ (2 Ci mol^{-1}), 10 mM $MgCl_2$, 50 mM Bicine at *p*H 8.22 or 8.30, and ~50 μg of protein were incubated for 10 min at 0, 0.26, 0.37, 0.62, 0.87 or 1.24 mM O_2 in a total volume of 0.5 ml. Reactions were initiated with 10 nmol of [1-^{3}H]RuBP (30 Ci mol^{-1}) and stopped after 5 min with 0.1 ml of a solution containing 0.5N H_2SO_4 and 0.05 M $ZnSO_4$. Labelled products of the carboxylase reaction ([^{3}H, ^{14}C]glycerate-P) and the oxygenase reaction (^{3}H-glycolate-P) were separated following glycolate-P hydrolysis and radioactivity in these compounds was quantified by scintillation spectroscopy. The ratio (ν_o/ν_c) of the two reactions, which is independent of RuBP concentration and enzyme activation state[9], was then plotted against the ratio of the O_2 and CO_2 concentrations present during assay. The specificity factor V_cK_o/V_oK_c is the reciprocal of the slope of this plot. Data is plotted for enzymes from *R. sphaeroides* (II) (■), *R. rubrum* (□), *A. flos-aquae* (●), *E. gracilis* (○), *S. obliquus* (▲) and *S. oleracea* (△).

still to be answered is whether further increases of the efficiency of the enzyme can be accomplished by the application of suitable selection pressure in combination with mutagenesis[7] or by other techniques.

Received 23 January; accepted 20 March 1981.

1. Ogren, W. L. & Bowes, G. *Nature new Biol.* **230,** 159–160 (1971).
2. Laing, W. A., Ogren, W. L. & Hageman, R. H. *Pl. Physiol.* **54,** 678–685 (1974).
3. Weissbach, A., Horecker, B. L. & Hurwitz, J. *J. biol. Chem.* **218,** 795–810 (1956).
4. Bowes, G., Ogren, W. L. & Hageman, R. H. *Biochem. biophys. Res. Commun.* **45,** 716–722 (1971).
5. Chollet, R. & Ogren, W. L. *Bot. Rev.* **41,** 137–179 (1975).
6. Ogren, W. L. in *Proc. 4th int. Congr. on Photosynthesis* (eds Hall, D. O., Coombs, J. & Goodwin, T. W.) 721–733 (The Biochemical Society, London, 1978).
7. Somerville, C. R. & Ogren, W. L. *Proc. natn. Acad. Sci. U.S.A.* **77,** 2684–2687 (1980).
8. Lorimer, G. H. & Andrews, T. J. *Nature* **243,** 359–360 (1973); *FEBS Lett.* **90,** 1–9 (1978).
9. Jordan, D. B. & Ogren, W. L. *Pl. Physiol.* **67,** 237–245 (1981).
10. Gibson, J. L. & Tabita, F. R. *J. biol. Chem.* **252,** 943–949 (1977).
11. Cloud, P. E. *Science* **160,** 729–736 (1968).
12. Badger, M. R., Kaplan, A. & Berry, J. A. *Yb. Carnegie Instn Wash.* **77,** 251–261 (1978).
13. Hatch, M. D. *Biochem. J.* **125,** 425–432 (1971).
14. Smith, B. N. & Robbins, M. J. in *Proc. 3rd int. Congr. on Photosynthesis* (ed. Avron, M.) 1579–1587 (Elsevier, Amsterdam, 1975).
15. Yeoh, H.-H., Badger, M. R. & Watson, L. *Pl. Physiol.* **66,** 1110–1112 (1980).

Modification of calmodulin function by enzymatic carboxyl methylation

Claude Gagnon*, Susan Kelly*, Vincent Manganiello†, Martha Vaughan†, Charles Odya†, Warren Strittmatter‡, Andrew Hoffman‡ & Fusao Hirata‡

*Department of Pharmacology, Faculty of Medicine,
Laval University, Quebec, Canada G1K 7P4
† National Heart, Lung and Blood Institute, Bethesda,
Maryland 20205, USA
‡ National Institute of Mental Health, Bethesda,
Maryland 20205, USA

Calmodulin is a ubiquitous calcium-binding protein that regulates a variety of enzymes such as adenylate cyclase, cyclic nucleotide phosphodiesterase, ATPase and protein kinase (for review see ref. 1). So far, no enzymatic modification of calmodulin has been shown to affect its function. Another ubiquitous protein, the enzyme protein-carboxyl methylase (PCM), modifies proteins post-translationally by methylating their free carboxyl groups and thus neutralizing negative charges[2–6]. This enzyme is one of three elements of the protein-carboxyl methylation system; the two others are the substrates, the methyl acceptor proteins (MAP) and the demethylating enzyme, protein methylesterase (PME)[7]. We report here that calmodulin is an exceptionally good substrate for PCM and that the enzymatic post-translational methylation of calmodulin inhibits its stimulatory effect on cyclic nucleotide phosphodiesterase. Furthermore, we present evidence that carboxyl methylation of calmodulin occurs in intact cells.

Bovine brain calmodulin was purified essentially according to Watterson *et al.*[8] with the exception that the acid–ammonium sulphate step was omitted. Methyl group (0.2 to 0.5 mol) was incorporated into calmodulin during 30 min at 37 °C in the following incubation medium: 0.5 μg calmodulin, 0.25 M sodium acetate *p*H 6.0, *S*-adenosyl-methionine 75 μM, 0.1% Triton X-100 and 2,000-fold purified PCM[9] in a final volume of 50 μl. The extent of methylation was determined in parallel samples by adding *S*-adenosyl-[*Me*-^{3}H]methionine to the incubation mixture, precipitating the calmodulin-^{3}H-methyl esters formed with 15% trichloroacetic acid, hydrolysing the ^{3}H-methyl esters with 1M sodium borate buffer, *p*H 11.0, and counting the radioactive methanol formed[9]. As a control, calmodulin or PCM or *S*-adenosyl-methionine was omitted from the incubation mixture. Calmodulin was an excellent substrate for PCM (Table 1). On a molar basis, methylation of calmodulin was 60–70% greater than that of luteinizing hormone, one of the best substrates known[10]. Polyacrylamide gel electrophoresis of methylated calmodulin in acid urea shows only one peak of radioactivity that co-migrates with calmodulin (Fig. 1).

Table 1 Comparative carboxyl-methylation of various purified proteins

Proteins	d.p.m. per pmol of protein
Calmodulin	1,140
Luteinizing hormone	684
Ovalbumin	201
Follicle-stimulating hormone	110
Catalase	143
Carboxylesterase	77
Bovine serum albumin	71
Aldolase	17
Haemoglobin	9
Cytochrome *c*	3
Myoglobin	3

About 1 μg of each protein was methylated as described elsewhere[9] using *S*-adenosyl-[*Me*-^{3}H]methionine as methyl donor. Protein concentration was determined according to Lowry *et al.*[16] using bovine serum albumin as a standard. Values are the mean of duplicate determinations in a single experiment, representative of three others.

To assess the effects of methylation on calmodulin function, samples were assayed using a calmodulin-free preparation of cyclic nucleotide phosphodiesterase. As protein-methyl esters are unstable at neutral and alkaline *p*H[11,12], the phosphodiesterase assay described by Kincaid *et al.*[13]. was performed at *p*H 6.0. In these conditions, the activation of cyclic GMP hydrolysis was similar to that observed at *p*H 7.5–8.0 with a 5–10-fold stimulation at optimum calmodulin and calcium concentrations.

Enzymatic carboxyl-methylation reduces the capacity of calmodulin to activate phosphodiesterase (Table 2) and the

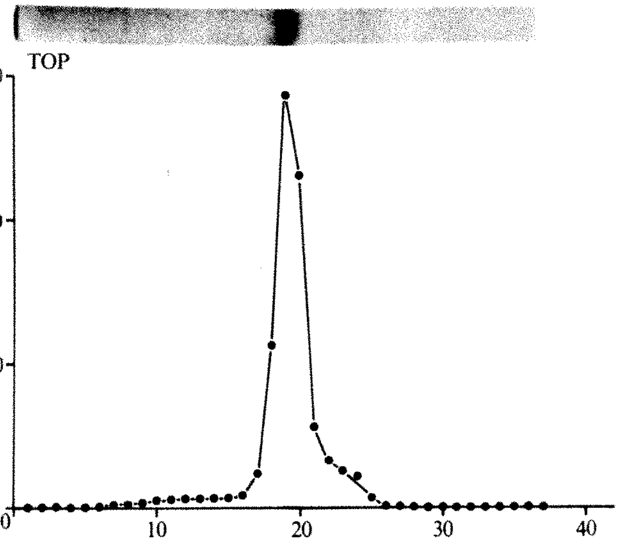


Fig. 1 Polyacrylamide gel electrophoresis of methylated calmodulin. Methylated calmodulin (10 μg) was subjected to electrophoresis on 10% polyacrylamide gel in an acetic acid–urea system as described elsewhere[12]. Gels were stained, destained, sliced and counted. The quantity of radioactivity present in the calmodulin band was identical to that recovered as [*Me*-^{3}H]methanol in a similar sample processed by the trichloroacetic acid method, which indicates that all radioactivity incorporated was in the form of methyl esters. Control gels (not shown), for example, protein-carboxyl methylase alone, reveal only one faint Coomassie blue-staining band at 20% of the gel length with no radioactivity detected in the whole gel.

degree of reduction correlates with the extent of methylation. Incubation of calmodulin in the methylating system from which PCM or *S*-adenosyl-methionine was omitted did not alter its capacity to activate phosphodiesterase (data not shown).

To determine whether the carboxyl methylation of calmodulin is simply an *in vitro* reaction or represents a biochemical reaction occurring *in vivo*, cells from two different cell lines were incubated with [*Me*-^{3}H]methionine. After a 3-h labelling period, a calmodulin-enriched fraction was isolated by phenothiazine–Sepharose affinity chromatography and analysed on acetic acid–urea gel. Two peaks of protein–^{3}H-methyl esters were observed on the electrophoretogram (Fig. 2). Whereas the identity of the first peak of radioactivity is unknown, the second, major peak corresponds to calmodulin as judged from the migration position of authentic calmodulin.

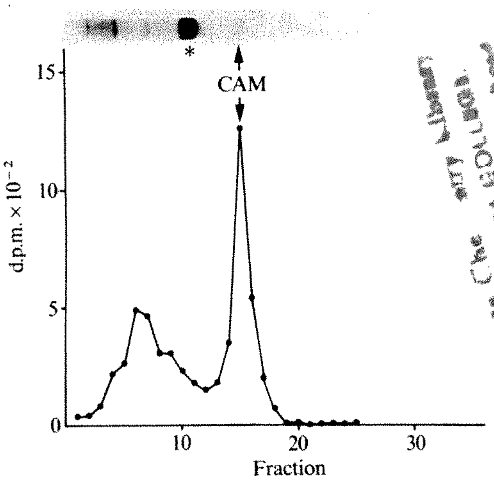


Fig. 2 Polyacrylamide gel electrophoresis of a calmodulin-enriched fraction isolated from intact cells labelled with [*Me*-^{3}H]methionine. Ascitic Walker carcinoma cells (1 g) in Earl's modified minimal essential medium supplemented with 10% fetal calf serum were labelled with 10 μCi ml^{-1} of [*Me*-^{3}H]methionine. After 3 h, cells were sonicated in 0.4 M sodium acetate buffer, *p*H 6.0, in the presence of 1 mM EGTA and a 100,000 *g* supernatant was prepared. This supernatant was adjusted to 5 mM $CaCl_2$ and loaded on to a 2-chloro-10 (3-aminopropyl)-phenothiazine–Sepharose column[18,19] equilibrated in acetate buffer and 5 mM $CaCl_2$. The column was washed in the equilibration buffer, followed by acetate buffer plus 1 mM $CaCl_2$. A calmodulin-enriched fraction was obtained by eluting the column with 10 mM EGTA in acetate buffer. As the chromatography was done in acidic conditions and small volumes of cells were used, only 50% of the original calmodulin (as judged by the phosphodiesterase assay) was recovered in the EGTA fraction. The other 50%, much contaminated by proteins, was recovered with 8 M urea. The EGTA fraction was precipitated with acetone at *p*H 2 or alternatively lyophilysed after acidification and analysed on acetic acid–urea gel[12]. After electrophoresis, the gels were stained, destained and sliced. Slices were incubated with 400 μl of sodium borate buffer, *p*H 11.0, for 6–12 h to allow the hydrolysis of protein-^{3}H-methyl esters to occur. The radioactive methanol formed was extracted using 6 ml of isoamyl alcohol–toluene (2:3 by volume). Two aliquots of 2.5 ml of the organic phase were collected; the radioactivity in one aliquot was counted directly whereas the second aliquot was evaporated to dryness before counting. Normally, over 95% of the radioactivity is volatile. Two peaks of radioactivity are seen on the gel and the second, major peak migrates to the same position as calmodulin. When purified calmodulin is methylated by purified PCM and processed through the phenothiazine–Sepharose column, the recovery of calmodulin-methyl esters in the EGTA plus urea fractions is ~45%. The instability of calmodulin-methyl esters is responsible for this low recovery. Very similar results were obtained with another cell line, the melanoma B16 F10 cells (data not shown). The arrows indicate the position of calmodulin and the asterisk indicates the position of an internal marker, cytochrome *c*.

Table 2 Effect of enzymatic carboxyl-methylation of calmodulin (CAM) on its capacity to stimulate cyclic nucleotide phosphodiesterase (PDE)

	PDE activation (c.p.m.)	CAM-activated PDE (c.p.m.)	% Inhibition of CAM activation
PDE alone	687		
PDE+CAM (15 ng ml^{-1})	1,250	563	
PDE+methyl-CAM (15 ng ml^{-1})	1,020	333	41
PDE+CAM (50 ng ml^{-1})	1,656	969	
PDE+methyl-CAM (50 ng ml^{-1})	1,142	455	53

After the methylation reaction, samples equivalent to 0.6 and 2 μl were assayed with a calmodulin-free preparation of cyclic nucleotide phosphodiesterase prepared according to Sharma and Wang[17]. Phosphodiesterase activity was assayed in a 400 μl assay containing 0.5 μM ^{3}H-labelled cyclic GMP, 5 mM $MgCl_2$, 0.1 mM $CaCl_2$, 0.2 M sodium acetate buffer, *p*H 6.0, 0.1 mg ml^{-1} ovalbumin, and native or methylated calmodulin (0.51 mol methyl group per mol calmodulin) at the indicated concentration. Mixtures were incubated at 30 °C for 15 min. Values are the mean of duplicated determinations in a single experiment, representative of three others.

These results provide the first evidence that carboxyl methylation of a protein substrate alters its function and of the regulation of calmodulin function as a result of covalent modification catalysed by a ubiquitous enzyme. Moreover, the data suggest that such regulation may be occurring in intact cells. Whether methylation alters the ability of calmodulin to bind calcium or to interact with the phosphodiesterase remains to be answered. Such answers will be difficult to obtain due to the instability of calmodulin–methyl esters at neutral and alkaline *p*H.

As the recently described mammalian protein methylesterase[9] can hydrolyse calmodulin-methyl esters (unpublished data) synthesized by PCM, the protein-carboxyl methylation system could, by its action on calmodulin, reversibly regulate the activity of the Ca^{2+}-dependent cyclic nucleotide phosphodiesterase as well as other enzymes activated by calmodulin.

In mammals, the protein-carboxyl methylation system seems to be involved in white blood cell chemotaxis and exocytotic secretion (for review, see ref. 7), two cellular processes in which calcium has a key role. Both PCM and protein methylesterase are activated after exposure of rabbit neutrophils to the synthetic chemoattractant fMet-Leu-Phe[14,15]. The data presented here suggest that calmodulin is a possible candidate for such regulation.

We acknowledge the support of the MRC of Canada, the Conseil de la Recherche en Santé du Québec and the Population Council.

Received 2 March; accepted 23 March 1981.

1. Cheung, W. Y. *Science* **207**, 19–27 (1980).
2. Liss, M., Maxam, A. M. & Cuprak, L. J. *J. biol. Chem.* **244**, 1617–1622 (1969).
3. Kim, S. & Paik, W. K. *J. biol. Chem.* **245**, 1806–1813 (1970).
4. Kim, S. & Paik, W. K. *Biochemistry* **10**, 3141–3145 (1971).
5. Morin, A. M. & Liss, M. *Biochem. biophys. Res. Commun.* **52**, 373–378 (1973).
6. Kim, S. *Archs Biochem. Biophys.* **157**, 476–484 (1973).
7. Gagnon, C. & Heisler, S. *Life Sci.* **25**, 993–1000 (1979).
8. Watterson, D. M., Harrilson, W. G. Jr, Keller, P. M., Sharief, F. & Vanaman, T. C. *J. biol. Chem.* **251**, 4501–4513 (1976).
9. Gagnon, C. *Biochem. biophys. Res. Commun.* **88**, 847–853 (1979).
10. Diliberto, E. J. Jr & Axelrod, J. *Proc. natn. Acad. Sci. U.S.A.* **71**, 1701–1704 (1974).
11. Kim, S. & Paik, W. K. *Experientia* **32**, 982–984 (1976).
12. Gagnon, C., Viveros, O. H., Diliberto, E. J. Jr & Axelrod, J. *J. biol Chem.* **353**, 3778–3781 (1978).
13. Kincaid, R. L., Manganiello, V. C. & Vaughan, M. *J. biol. Chem.* **254**, 4970–4973 (1979).
14. O'Dea, R. F. *et al. Nature* **272**, 462–464 (1978).
15. Venkatasubramanian, K. *et al. Molec. Immun.* **17**, 201–207 (1980).
16. Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193**, 265–275 (1951).
17. Sharma, R. K. & Wang, J. H. in *Advances in Cyclic Nucleotide Research* (eds Brooker, G., Greengard, P. & Robinson, G. A.) 187–198 (Raven, New York, 1979).
18. Jamieson, G. A. & Vanaman, T. C. *Biochem. biophys. Res. Commun.* **90**, 1048–1056 (1979).
19. Van Eldik, L. J., Grossman, A. R., Iverson, D. B. & Watterson, D. M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1912–1916 (1980).

BOOK REVIEWS

Probity in science: the case of Cyril Burt

J.M. Thoday

A Balance Sheet on Burt. Supplement to the *Bulletin of The British Psychological Society,* Vol.33. Edited by Halla Beloff. ISBN 0-901-71513-1. Pp.38. (British Psychological Society: 1980.) £1.

For most of the life of the subject of genetics, and longer, students of "heredity" have been faced with two controversies, each of which has had political undertones, and sometimes overtones. One concerns the inheritance of acquired characters, the other the inheritance of variation in ability. Debate about the inheritance of acquired characteristics has had its scandals and its tragedies, and is not wholly dead.

The other controversy is very much alive today. And this pamphlet concerns one of its major scandals, for it reports the results of a meeting of the British Psychological Society, called to consider the implications of the evidence that the contributions of Sir Cyril Burt to psychometrics were at least in part based on fabricated data.

Well before this meeting took place, it had become clear that some of Burt's data were suspect if not fabricated. When Kamin first reported evidence of this, reactions seem to have ranged from delight to shock, to sheer disbelief and sometimes anger. Those who reacted in anger (and the Society itself) have of course been accused of a cover-up, but it takes time for a profession to come round to the view that one of its most respected members, on whose data they have relied, was fraudulent.

Be that as it may, most have now come round to the view that all of Burt's data must be treated as suspect and the view is confirmed by Hearnshaw's sympathetic biography (*Cyril Burt: Psychologist* Hodder & Stoughton, 1979; reviewed in *Nature* **282**, 150), and in this pamphlet. What remains is to ask how this sort of thing can be prevented in future and how deleting Burt's data affects our answers to the relevant questions about IQ.

I think the clear general message of this pamphlet and of the whole story is that editors of professional journals should be more inclined to publish basic data, and require careful accounts of how raw data have been treated; in addition, as it suggests, editors should be required to submit their own papers to external referees, nominated by someone other than the editor, for many of Burt's publications were not refereed. For the rest we must be grateful to those who, whatever their motives, did the critical research that has shown us Burt's data are not to be relied on. We need vigilance if such questionable data are to be exposed. Mainly it must depend upon the vigilance of the referees and experts. But I think it is at this level that ideological motivation may have a place in science. If strong distaste for scientific conclusions leads to honest and intelligent criticism of the data base or logical base for those conclusions, this is to the good. The trouble is, however, that such criticism tends to be self-defeating, for it is often presented as if the motivation was itself a justification and thus rouses its own antagonists. Hence as the IQ controversy exemplifies, the criticism simply polarizes two antithetical camps, rather than helping to get us closer to the truth.

To turn to what I think is the more important matter: how does the rejection of Burt's data leave the field? That is to say, how do we answer the following questions on the basis only of other data?

(1) *Is relative performance in IQ tests affected by genetic variance, environmental variance, or both, and if both, how much of each?*

I think it is only the last part of this question that is affected by the deletion of Burt's data. His heritability estimates are higher than most. But most estimates, whatever kind of relatives are compared, give correlations between relatives that imply reasonably high heritability, and all imply that heritability is incomplete, as do also adoption studies. This latter has never been in doubt: I do not think anyone has seriously maintained that IQ variance is all genetic, though Burt wished to define "intelligence" as the genetic component.

Of course much will have to be recalculated, and good deal of that which depended heavily on Burt's data will need to be rewritten. But Rowe and Plomin (*Behaviour Genetics* **8**, 81; 1978) and Paul (*Behaviour Genetics* **10**, 277; 1980) have compared Burt's data with those of other workers and, though Burt's correlations are rather more extreme, only one differs significantly.

We are, therefore, still left with evidence for rather high heritability, in the face of which the only arguments a determined environmentalist can use are to suppose that all the data are spurious, or to rest his case on the effects of environmental correlation between relatives mimicking completely the theoretical genetic correlations. Such environmental correlations must be a difficulty of human biometrical genetics, but it seems hard to believe that they could explain all the evidence, especially the evidence that the correlation between monozygotic twins reared apart (0.74) is so much higher than dizygotic twins (0.53) or sibs (0.54) reared together.

Such correlations, when assortative mating is allowed for, imply rather high heritability, which, however, is an upper limit, as some part of the correlation will be environmental. But I doubt if the data could be made consistent with a heritability of less than say 0.4, and point out that all the significant "political" issues with respect to within-group comparisons arise qualitatively whether heritability is 0.2, 0.5 or 0.8, and also agree with Clarke's stress that we do not expect different groups to give the same heritabilities.

Camera Press

Cyril Burt — what are the consequences of the exposure of his fabricated data?

(2) *What, if any, are the differences or mean IQ between groups (social or racial), and what environmental or genetic components have such differences?*

The removal of Burt's data still leaves the evidence that groups differ. I have spelt out elsewhere the problems of determining the causes of such differences and the importance of the fact that group distributions overlap, so that stress on the differences of mean is unfair to individuals. This is true whatever the genetic component. Of course it is called scientific racism to have an open mind on the nature–nurture controversy, but I would repeat (see *J. Biol. Educ.* **6**, 323; 1972) that there are grave dangers in basing policies such as affirmative action on the assumption that all differences arise from environmental causes and all individuals are of identical potential.

Amongst the matters unaffected by rejecting Burt's data remains the evidence that IQ and social mobility are correlated. Given significant heritability of IQ this implies that part of class differences in mean must become genetic if they are not already so. It is not reasonable to ignore such evidence, whatever one's ideology. I have also pointed out (*J. Biosoc. Sci.* Suppl. **1**, 3; 1969) that when group differences involve a morphological marker, as with skin colour, it is at present formally impossible to determine whether any difference of mean distinguishing the groups is all environmental, all genetic or anything in between. But group mean differences would hardly matter if we treated individuals as such, except if correlations between fertility and IQ differed between groups which, as Jensen pointed out, might occur. If so, heritability within the group becomes important as the predictor of future change, and of future mean differences.

(3) *Do IQ tests measure something whose variation matters?*

Some (see, for example, Rose *Symp. Inst. Biol.* **22**, 191; 1975) do not think that IQ tests measure anything.

However, these tests rank individuals on a scale just as stature measures do, and differ from stature measures in only three ways: the measurement error may be larger, IQ scales have no zero, and stature measures can be applied equally accurately, or inaccurately, to doors and trousers. But errors in both measurements can be assessed, and the zero point is irrelevant to biometrical analyses since these are solely concerned with deviations from means. The real difference concerns measuring doors and this is what leads many to think IQ tests only measure ability to do IQ tests. But, though we cannot measure educational doors with such tests, there is significant though incomplete correlation with future educational performance, and various other aspects of "success" in life as Terman's investigations first demonstrated. If, therefore, we consider that the attributes of which IQ tests are partial predictors matter, either to the individual or society, then they do measure something that matters. There might be better measures, complementary measures might make more powerful combined predictors. But they do seem to work, and assessment of this is not affected by the deletion of Burt's data. I therefore tend to disagree with Blackman, a contributor to the pamphlet, whose suggested investigations concerning individual development seem to ignore that the essential issues concern the nature and causes of between-individual variance.

(4) *Should IQ tests, complemented or not, be used in any way?*

This, of course, is the ideologically loaded question and it is not to be answered by a scientist as scientist but only as citizen. Whether they should be used for investigative purposes depends on your belief in knowledge; whether for selective purposes perhaps upon your ideology.

But I would point out that there is an element of fashion in ideology. Dr Gillie says here that "a generation or more of children may have suffered" as a result partly of Burt's efforts. I can only assure him that the editor of this pamphlet is right about the different view that was taken of the 1944 British Education Act in its early days. Many of my generation when young regarded it as a great socially progressive enactment!

I conclude that Burt's exposure makes very little difference to our knowledge of IQ and its heritability and agree with Anne Clarke when she says in her essay that "I believe that now that Burt's embarrassing results have been disposed of, we can get on with building a solid science of human differences".

In the short run self-criticism is the best insurance against error. Mutual criticism leading to repeat investigation will expose error in the long run, but in an area where conclusions may affect political action, the long run can be too long. In such areas it is particularly important to avoid antagonizing opponents, but to write as carefully and objectively as possible and, if there are data to present, to present them as they are so that others can test them critically. □

J.M. Thoday is Arthur Balfour Professor and Head of Department in the Department of Genetics, University of Cambridge.

Solar puzzles for radio astronomers

David Layzer

Radio Physics of the Sun. Edited by M.R. Kundu and T.E. Gergely. Pp.475. ISBN hbk 90-277-1120-8; ISBN pbk 90-277-1121-6. (Reidel: 1980.) Hbk Dfl.115, $60.50; pbk Dfl.55, $28.95.

THESE well-edited proceedings of an international conference held in August 1979 and sponsored by the International Astronomical Union contain several excellent short reviews, over 50 brief research reports and an edited transcript of the discussion following each communication. Most of the reviews and reports deal with observations of solar radio emission and with the underlying emission mechanisms; a few discuss the radio observations in a broader observational and theoretical context.

In recent years much new information about the quiet Sun, active regions and solar bursts has come from high-resolution microwave observations with large arrays and from observations at metre wavelength with the Culgoora imaging radio telescope in Australia. High-resolution microwave measurements of circular polarization now delineate small-scale magnetic fields in the lower corona. Heliograms at metre wavelengths show the hole (open field-line) and arch (closed field-line) structure so conspicuous in soft X-ray pictures. Sequences of metre-wavelength heliograms have enabled radio astronomers to track moving Type IV solar bursts and measure their polarization, and have shown that the physics of these transients is not as well understood as had been supposed.

Theories of Type III bursts are the subject of a review paper and several brief reports. Certain aspects of the disturbances responsible for these bursts — in particular, how they maintain their identity for so long (tens of seconds) and over such long path lengths (millions of kilometres) — have puzzled theoreticians for nearly 20 years. At this conference, two distinct and mutually inconsistent approaches to these problems are described, one based on quasilinear theory, the other on strong-turbulence theory. Unfortunately, the two approaches are not given an adequate opportunity to confront each other in the recorded discussion.

This volume (and presumably the conference itself) would have been greatly improved by the inclusion of a few introductory and summary reports of sufficient breadth and depth to provide context and continuity. The invited reviews that open each session are excellent, but too short and too limited in scope; and there are no summaries at all. The volume is dedicated to Stefan Smerd. One misses the breadth and depth of the two classic reviews of solar radio astronomy, both published in *Annual Review of Astronomy and Astrophysics*, that he wrote in collaboration with Paul Wild and A.A. Weiss ("Solar Bursts" by J.P. Wild, S.F. Smerd and A.A. Weiss **1**, 291–366, 1963; "Radio Bursts from the Solar Corona" by J.P. Wild and S.F. Smerd **10**, 159–196, 1972). □

David Layzer is Donald H. Menzel Professor of Astrophysics at Harvard University.

A matter of some gravity

W. D. Allen

The Search for Gravity Waves. By P.C.W. Davies. Pp.144. ISBN 0-521-23197-3. (Cambridge University Press: 1980.) £5.95, $13.95.

THERE is no doubt that Joseph Weber's claim in 1969, to have established evidence for gravitational radiation, has, even if unsubstantiated, stimulated great interest. Thus, most reports of the unsuccessful attempts to corroborate Weber have attracted notice in the popular press. Now, Professor Davies's book brings the subject within the range of the intelligent but non-specialist reader. The book has all the hallmarks of a good exposé. There is very little algebra; the author is concerned throughout to make basic principles picturable, and any formulae derived are brought firmly to earth with order-of-magnitude estimates. Where the theory is complex or obscure, Davies directs attention to the variables involved, deduces the form of the equation linking the variables on dimensional grounds and then supplies the constant of proportionality from the theoretical formulae which results from detailed analysis.

Also, the book develops a logical thread of argument throughout, beginning naturally with electromagnetic radiation. This radiation has many parallels with gravitational radiation, and is familiar in everyday life; but there are important differences, which require emphasis. This opening chapter, on electromagnetic radiation, is strongly to be recommended to those who think the subject can be understood only by reference to vector algebra and Maxwell's equations. Next follows the counterpart chapter on gravity waves, which is the core of the book. In particular, considerable space is given to making clear the role of second-order variations in a gravitational field, known to us all in the phenomenon of tides. For gravity waves, both generation and detection depend on these tidal forces. Next, principles become practice. Formulae are developed which show that the radiation power flux density (in SI units) is negligible for any terrestrial gravity wave generator; and is also negligible, at the Earth's surface, for binary stars, the astronomical gravity-wave generators whose output can be accurately calculated. Two other astronomical candidates are discussed; the vibrating neutron star, relict of a supernova explosion, and the collision of two black holes, each of stellar mass magnitude. Given the right distance from the Earth, i.e. the appropriate degree of optimism, and accepting the necessary approximations in the calculations, the terrestrial radiation flux is appreciable. Having established possible terrestrial fluxes, the author turns to detectors. The cross-section for detection of a tonne of aluminium is calculated, and it is shown that, taking the radiation flux figures given previously, detection is possible.

Finally, the chapter "Have they been seen?" is mainly on Weber's studies, with a brief glance at another experimental methods, and on a description of observations on the slowing down of a binary pulsar. The decay rate of the period closely matches prediction, although alternative explanations need examination.

There are so many points of interest in the book that comment is somewhat random. First, the early disbelief in gravity waves is rather glossed over, perhaps rightly. But one would still like to believe in the authenticity of a remark attributed to a distinguished astronomer in the 1920s: "Gravity waves travel with the speed of thought". Second, the survey of experiments in terrestrial detection is disappointing; there are two experimental groups still active in Britain and available for advice, and a better survey could have been made, even at the expense of a fractional increase in the book's size.

Weber said in the late 1960s that "In the last 15 years, I have learnt much more about Einstein's theory of Brownian motion that I have about Einstein's theory of gravity". This exactly sums up the field of the gravity wave experiments. If, in a future edition, Professor Davies writes as clearly about Brownian motion as he does about radiation, then, irrespective of gravity waves, he will have written an excellent undergraduate text.

I hope this "popular" book has popular acclaim. ☐

W. D. Allen is Visiting Professor in the Department of Electrostatics, University of Southampton.

Lactation reference

Alan S. McNeilly

Hormonal Control of Lactation. By A.T. Cowie, Isabel A. Forsyth and I.C. Hart. Pp.275. ISBN 3-540-09680-9. (Springer-Verlag: 1980.) DM88, $52.

CONSIDERING the essential role of lactation in the reproductive cycle of mammals, it is suprisingly often neglected in the teaching of reproductive biology. This may in part be due to the absence of a book which reviews the present state of knowledge on the hormonal control of lactation in a comprehensive but straightforward manner. In the past, texts have been either austere tomes or short monographs essentially in the form of notes for students. I believe that the book by Cowie, Forsyth and Hart has now bridged this gap. It is well conceived in leading the reader

NEW WORKS IN THE LIFE SCIENCES FROM GORDON AND BREACH...

This year we celebrate twenty years of publishing scientific books and journals Among our recent works in the life sciences:

A New Journal

COMMENTS ON MOLECULAR AND CELLULAR BIOPHYSICS
Burt V. Bronk, Clemson University, and J.W. Longworth, Oak Ridge National Laboratory, USA, Editors

Launched last December, this journal is the first in a planned series of **Comments on Modern Biology** to complement the well-established G & B **Comments on Modern Physics.** It is designed for the quantitatively oriented molecular biologist, biochemist and cell biologist as well as for the physicist interested in biology.
6 issues per volume $99.00 £42.00
For a free sample, please write to our Customer Service.

HEPARIN (AND RELATED POLYSACCHARIDES):
Structure and Functional Properties
Wayne D. Comper, Monash University, Australia

Here is Volume 7 in a series of monographs on polymers and their applications. Authoritative and up-to-date, each is written by an expert in his field. HEPARIN covers, in brief, **monosaccharide sequence; degradation; secondary and higher order structures; function; tissue distribution and cellular localisation; and extracellular interaction.**
1981 262pp approx. $45.00 £24.00

A THEORETICAL APPROACH TO THE PRESELECTION OF CARCINOGENS AND CHEMICAL CARCINOGENESIS
Veljko Veljkovic, Boris Kidric Institute of Nuclear Sciences, Yugoslavia

The author presents a simple valence theory of carcinogenesis, using a new approach to quantitative treatment of molecular biology and problems of chemical carcinogenesis. Based on theoretical results, a new quick test for predicting carcinogenicity of chemical substances is given, together with guidance to the selection, design and use of drugs in therapy.
1981 124pp $36.50 £19.21

Would you like to receive our NEW BOOKS LIST and our forthcoming JOURNALS CATALOGUE? For free copies please write to Customer Service. Prices may be subject to change

G+B

Gordon and Breach
Science Publishers
42 William IV Street
London WC2N 4DE

Circle No.30 on Reader Enquiry Card.

through the chemistry of the hormones involved in lactation to chapters on the growth and development of the mammary gland, hormonal control of milk secretion and milk ejection, assays for the hormones and a section on hormonal receptors and their mechanisms. In these it is possible to gain a comparative picture, as far as our knowledge will allow, of the hormonal requirements for mammary development in species as diverse as the Californian leaf-nosed bat, rabbit and monkey. Indeed, the book emphasizes how limited is our knowledge of the endocrine events controlling not only lactation but reproduction as a whole in the vast majority of mammals.

It is satisfying to see that the authors have included the human being in their considerations. Too often in the past has knowledge gained from animals been ignored or remained unknown in clinical practice, as is clearly illustrated in the section on milk ejection. Some 5,000 years ago milkers were aware that an agitated cow did not give milk readily. Similarly, Cooper in 1840 stated that successful breast-feeding also needs a calm and tranquil mind, a point which many busy post-natal wards in the modern hospital would do well to remember.

It is clear that many of the recent advances in the field covered by this book have resulted from detailed studies using radioimmunoassays of the hormones involved. As the authors discuss, it is salutary to remember that these methods do not necessarily measure the biological activity of these hormones: will our ideas again need changing when true bioactivity is measured?

A minor criticism is the inclusion of many illustrations from the authors' own studies when, in some instances, others would have been more informative. Nevertheless I found this book easy to read, full of invaluable information and in parts enjoyable; for example, one of the authors describes the time when, having previously worked with milligram quantities, he received a bottle containing half a kilogram of crystalline progesterone for studies on the induction of lactation.

This monograph is well presented and should become a standard reference for all those interested in lactation. ☐

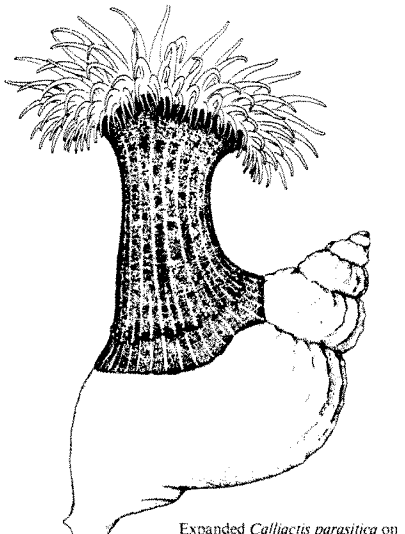
Alan S. McNeilly is a Research Scientist in the MRC Unit of Reproductive Biology, Edinburgh.

Anthozoan beauty

C.M. Yonge

British Anthozoa. Synopses of the British Fauna, 18. By R.L. Manuel. Pp.241. ISBN 0-12-470050-0. (Academic: 1981.) £7.50, $18.

THIS is the latest and, next but one, the largest of the *Synopses of the British Fauna*. The series is admirably edited by Doris Kermack on behalf of the Linnean Society in cooperation, in recent parts, with R.S.K. Barnes of the Estuarine and Brackish-water Sciences Association.



Expanded *Calliactis parasitica* on the shell of *Buccinum undulatum*.

While the Anthozoa, in the form of stony, horny and soft corals, flourish most abundantly in shallow tropical waters, the British fauna comprises no less than 73 species. The greatest number are sea anemones, among them six species of the closely related ahermatypic (non-reef-building) corals, three cerianthids (one only known in the larval state) and eight zoanthids. The octocorallines are represented by alcyonarians such as the well-known dead-man's fingers (*Alcyonium digitatum* L.), together with two gorgonids (rock-attached sea fans) and four pennatulids (sea pens) that live unattached but erect in bottom muds.

Most anthozoans, particularly the anemones, are animals of great beauty of colour and delicate design. They have to be described when alive and fully expanded. An extremely high standard of illustration was set by Philip Henry Gosse in *Actinologia Britannica* (1860) and by Alan Stephenson in the two volumes of his Ray Society *British Sea Anemones* (1928, 1935). This book has a wider coverage and, although inevitably denied colour and with only two pages of photographs, the line drawings, largely of living animals and showing pattern and texture, are so beautifully executed as fully to maintain this standard. This is an altogether admirable guide. ☐

Sir Maurice Yonge is an Honorary Fellow in Zoology at the University of Edinburgh.

What's new in hybridoma technology?

Carlo M. Croce

Monoclonal Antibodies. Hybridomas: A New Dimension in Biological Analyses. Edited by R.H. Kennett, T.J. McKearn and K.B. Bechtol. Pp.423. ISBN 0-306-40408-7. (Plenum Press: 1980.) $22.50, £18.59.

THE possibility of producing hybridomas secreting monoclonal antibodies of predefined specificities has changed the field of immunology considerably and has provided powerful reagents for biomedical investigation. This book is an attempt to summarize the most recent developments in this area of research.

The book is divided into two major sections. The first concerns the production of hybridomas and the many possible applications of hybridoma technology, for example in studies of the structure and the genetics of immunoglobulins, in the study of the cell surface, in the definition of antigens specific for cells of different lineages and for cells at different stages of differentiation, and in the study of human tumour-associated antigens. Considerable emphasis is given to the analysis of the cell surface by the use of monoclonal antibodies. The second and much smaller section is an appendix cataloguing various methods for the production and the characterization of monoclonal antibodies. A short summary of every technique related to the production, storage and characterization of hybridoma clones is also included.

This book will be of considerable help to investigators who are not familiar with tissue culture and somatic cell fusion techniques and who intend to produce hybridomas secreting monoclonal antibodies for a variety of purposes. It also provides a good overview of some of the most recent applications of hybridoma technology, particularly in the study of the cell surfaces of malignant cells and of cells of different lineages or stages of differentiation. Although the articles are, in general, well presented and informative, they are rather repetitious in that similar studies carried out in different laboratories are described. It would have been much more useful, I believe, if the editors had included only those articles that reflect the "state of the art" in the areas of application of the monoclonal antibodies to biomedical research. They might also have added an overview of some of the more important contributions in this area.

Overall, this volume provides the necessary information to carry out experiments involving monoclonal antibodies and should be an asset to investigators who wish to become familiar with hybridoma technology. ☐

Carlo M. Croce is Associate Director of The Wistar Institute, Philadelphia.

BOOKS RECEIVED

Physics

BILLINGTON, E. W. and TATE, A. The Physics of Deformation and Flow. Pp.626. ISBN 0-07-005285-9. (McGraw -Hill: 1981.) $59.

CHEN, R. and KIRSH, Y. Analysis of Thermally Stimulated Processes. International Series on the Science of the Solid State, Vol.15. Pp.361. ISBN 0-08-022930-1. (Pergamon: 1981.) £25, $60.

FAVA, R. A. (ed.). Polymers. Part C: Physical Properties. Methods of Experimental Physics, Vol.16. Pp.528. ISBN 0-12-475958-0. (Academic: 1980.) £33.20, $59.

HOYLE, F. and NARLIKAR, J. The Physics — Astronomy Frontier. Pp.483. ISBN 0-7167-1160-5. (W. H. Freeman: 1981.) £10.95.

JAMES, G. D. *et al.* Nuclear Fission and Neutron-Induced Fission Cross-Sections. Neutron Physics and Nuclear Data in Science and Technology, Vol.1. Pp.277. ISBN 0-08-026125-6. (Pergamon: 1981.) £21, $50.

KHALATNIKOV, I. M. (ed.) Soviet Scientific Reviews, Section A. Physics Reviews, Vol.2 (1980). Pp.484. ISBN 3-7186-0017-X. (Harwood Academic Publishers, New York: 1980.) $98.

KING, R. W. P., SMITH, G. S., OWENS, M. and WU, T. T. Antennas in Matter. Fundamentals, Theory and Applications. Pp.868. ISBN 0-262-11074-1. (MIT Press: 1981.) $75.

LIPSON, S. G. and LIPSON, H. Optical Physics. 2nd Edn. Pp.463. Hbk ISBN 0-521-22630-9; pbk ISBN 0-521-29584-X (Cambridge University Press: 1981.) Hbk £25; pbk £9.95.

RUMER, Yu. B. and RYVKIN, M. Sh. Thermodynamics, Statistical Physics, and Kinetics. Pp.600. ISBN 0-8285-1853-X. (Mir, Moscow/Imported Publications, Chicago: 1980.) 12.00.

TRAN THANH VAN, J. (ed.). Electroweak Interactions and Unified Theories, Vol.II. Proceedings of the 15th Rencontre De Moriond, Les Arcs-Savoie, France, March 1980. Pp.605. ISBN 2-86332-007-6. (Editions Frontieres, DERUX: 1980.) Np.

WELFORD, W. T. Optics. 2nd Edn. Pp.150. Hbk ISBN 0-19-851846-3; pbk ISBN 0-19-851847-1. (Oxford University Press: 1981.) Hbk £12.50; pbk £5.95.

WILKINSON, E. (ed.). Progress in Particle and Nuclear Physics, Vol.5. Pp.285. ISBN 0-08-027109-X. (Pergamon: 1981.) Np.

Chemistry

CHANG, R. Physical Chemistry with Applications to Biological Systems. 2nd Edn. Pp.659. ISBN 0-02-321040-0. (Collier Macmillan: 1981.) £12.95.

CONWAY, B. E. Ionic Hydration in Chemistry and Biophysics. Studies in Physical and Theoretical Chemistry, 12. Pp.768. ISBN 0-444-41947-0. (Elsevier Scientific: 1980.) $131.75, Dfl. 270.

HAY, R. K. M. Chemistry for Agriculture and Ecology. A Foundation Course. Pp.243. Pbk ISBN 0-632-00699-4. (Blackwell Scientific: 1981.) Pbk £5.50.

HOLMES, R. R. Pentacoordinated Phosphorus. Vol.I, Structure and Spectroscopy. ACS Monograph, 175. Pp.479. ISBN 0-8412-0458-6. (American Chemical Society: 1980.) $92.

RATAJCZAK, H. and ORVILLE-THOMAS, W. J. (eds). Molecular Interactions, Vol.2. Pp.627. ISBN 0-471-27681-2. (Wiley: 1981.) £39, $110.

Biological Sciences

BOREK, C. and WILLIAMS, G. M. (eds). Differentiation and Carcinogenesis in Liver Cell Cultures. Annals of the New York Academy of Sciences, Vol.349. Pp.429. Hbk ISBN 0-89766-087-0; pbk ISBN 0-89766-088-9. (New York Academy of Sciences, New York: 1980.) Hbk $83; pbk np.

BOURNE, G. H. (ed.). Hearts and Heart-Like Organs. Vol.3, Pathology and Surgery of the Heart. Pp.431. ISBN 0-12-119403-5. (Academic: 1980.) $53.

BOURNE, G. H. and DANIELLI, J. F. (eds). International Review of Cytology, Vol.69. Pp.339. ISBN 0-12-364469-0. (Academic: 1981.) Np.

BORGAONKAR, D. S. Chromosomal Variation in Man. A Catalogue of Chromosomal Varieties and Anomalies. 3rd Edn. Pp.846. ISBN 0-8451-0206-0. (Alan R. Liss, New York: 1981.) $58.

BROWDER, L. W. Developmental Biology. Pp.602. ISBN 0-03-056748-3. (Saunders College, Philadelphia: 1980.) Np.

BROWN, F. *et al.* (eds). An Arctic Ecosystem: The Coastal Tundra at Barrow, Alaska. US/IBP Synthesis Series, 12. Pp.571. ISBN 0-87933-370-7. (Academic: 1981.) $29.50.

BROWN, V. The Amateur Naturalist's Handbook. Pp.419. Hbk ISBN 0-13-023739-6; pbk ISBN 0-13-023721-3. (Prentice/Hall International: 1981.) Hbk £10.35; pbk £4.50.

BÜLBRING, E. *et al.* (eds). Smooth Muscle: An Assessment of Current Knowledge. Pp.563. ISBN 0-7131-4348-7. (Edward Arnold, London: 1981.) £45.

HENDRY, B. Membrane Physiology and Cell Excitation. Pp.159. Hbk ISBN 0-7099-0148-8; pbk ISBN 0-7099-0149-6. (Croom Helm, London: 1981.) Hbk £9.95; pbk Np.

HIRAYAMA, T., WATERHOUSE, J.A.H. and FRAUMENI, J.F. Jr (eds). Cancer Risks by Site. UICC Technical Report Series, Vol.41. Pp.217. Flexi ISBN 92-9018-041-2. (International Union Against Cancer (UICC), Geneva: 1980.) SwFr.20.

INCHLEY, C.J. Immunobiology. The Institute of Biology's Studies in Biology, No.128. Pp.82. Flexi ISBN 0-7131-2808-9. (Edward Arnold, London: 1981.) £2.40.

JONES, J.C. The Anatomy of the Grasshopper (*Romalea microptera*) Pp.281. (loose leaf) ISBN 0-398-04126-1. (Charles C. Thomas, Springfield, Illinois: 1981.) $27.50.

KNIPE, A.C. and WATTS, W.E. Organic Reaction Mechanisms, 1979. An Annual Survey Covering the Literature dated December, 1978 through November 1979. Pp.759. ISBN 0-471-27818. (Wiley: 1981.) Np.

KRAUSE, H. The Mammoth—In Ice and Snow? Cold-Adaptation of Woolly Mammoth: Fact or Fiction? Pp.107. (Self-Publisher, 7 Stuttgart 40, Kameralamtsstr.22.) Flexi DM 21.30.

LAVERACK, M.S. and COSENS, D.J. (eds). Sense Organs. Pp.394. ISBN 0-216-91094-3. (Blackie: 1981.) £29.75.

LEVY, G. and AMITAI, P. Scorpiones. Fauna Palaestina I. Pp.134. ISBN 965-208-016-0. (The Israel Academy of Sciences and Humanities, Jerusalem: 1980.) $20.

LEWIS, M. and ROSENBLUM, L.A. (eds). The Uncommon Child. Genesis of Behavior, Vol.3. Pp.342. ISBN 0-306-40499-0. (Plenum: 1981.) $25.

LUCKWILL, L.C. Growth Regulators in Crop Production. The Institute of Biology's Studies in Biology, No.129. Pp.59. Flexi ISBN 0-7131-2816-X. (Edward Arnold, London: 1981.) £1.95.

McLAREN, D.S. A Colour Atlas of Nutritional Disorders. Pp.109. ISBN 0-7234-0757-6. (Wolfe Medical Publications, London: 1981.) £15.

McLUSKY, D.S. The Estuarine Ecosystem. Pp.150. Hbk ISBN 0-216-91115-X; pbk ISBN 0-216-91116-8. (Blackie: 1981.) Hbk £12.75; pbk £6.25.

MOORE, K.L. Clinically Oriented Anatomy. Pp.1257. ISBN 0-683-06146-1. (Williams & Wilkins, Baltimore/London: 1980.) Np.

OLNEY, P.J.S. (ed.). 1980 International Zoo Yearbook, Vol.20. Pp.521. ISBN 0074-9664. (The Zoological Society of London: 1980.) Hbk £21; pbk £18.

PHILLIPS, M. and BAETZ, A. (eds). Diet and Resistance to Disease. Advances in Experimental Medicine and Biology, Vol.135. Proceedings of the Symposium held at the American Chemical Society Agricultural and Food Division Meeting, March 1980, Houston, Texas. Pp.220. ISBN 0-306-40636-5. (Plenum: 1981.) Np.

PUECH, P. and KREBS, R. (eds.) 4th International Adalat Symposium. New Therapy of Ischemic Heart Disease. Proceedings of the Symposium held at Paris, October 1979. Pp.298. ISBN 90-219-9440-2. (Excerpta Medica, Amsterdam: 1980.) $58.50, Dfl. 120.

RANDALL, D.J. *et al.* The Evolution of Air Breathing in Vertebrates. Pp.133. ISBN 0-521-22259-1. (Cambridge University Press: 1981.) £15.

RATTAZI, M.C., SCANDALIOS, J.G. and WHITT, G.S. (eds). Isozymes. Current Topics in Biological and Medical Research, Vol.4. Pp.218. ISBN 0-8451-0253-2. (Alan R. Liss, New York: 1980.) $26, Dfl.78.

REICHLE, D.E. (ed.). Dynamic Properties of Forest Ecosystems. International Biological Programme, 23. Pp.683. ISBN 0-521-22508-6. (Cambridge University Press: 1981.) £50.

ROLINSON, G.N. and WATSON, A. (eds). Augmentin Clavulanate-Potentiated Amoxycillin. Proceedings of the 1st Symposium, July 1980. Pp.308. ISBN 90-219-9487-9. (Excerpta Medica, Amsterdam: 1980.) $56, Dfl.115.

SAFAI, B. and GOOD, R.A. (eds). Immunodermatology. Comprehensive Immunology, Vol.7. Pp.717. ISBN 0-306-40380-3. (Plenum: 1981.) $49.50.

SAKAMOTO, S., TOJO, S. and NAKAYAMA, T. (eds). Gynecology and Obstetrics. Internatinal Congress Series, No.512. Proceedings of the IX World Congress, Tokyo, October 1979. Pp. 1343. ISBN 90-219-0440-3. (Excerpta Medica, Amsterdam: 1980.) $219, Dfl.450.

SALMON, S.E. (ed.). Cloning of Human Tumor Stem Cells. Progress in Clinical and Biological Research, Vol.48. Pp.366. ISBN 0-8451-0048-3. (Alan R. Liss, New York: 1980.) $44, Dfl.132.

SCHEFFER, V.B. A Natural History of Marine Mammals. Pp.157. Pbk ISBN 0-684-16952-5. (Charles Scribner's: 1981.) Pbk $5.95.

SCHMAHL, D. Maligne Tumoren. Entstehung, Wachstum, Chemotherapie. Pp.572. ISBN 3-87193-0628. (Editio Cantor Aulendorf: 1981.) Np.

SHAMOO, A.E. 2nd International Conference on Carriers and Channels in Biological Systems-Transport Proteins. Annals of the New York Academy of Sciences, Vol.358. Pp.387. Hbk ISBN 0-89766-105-2; pbk ISBN 0-89766-106-0. (The New York Academy of Sciences, New York: 1980.) Hbk $75.

SHORT, R.V. and WEIR, B.J. (eds). The Great Apes of Africa. Journal of Reproduction and Fertility Supplement, No.28. Proceedings of a Symposium held at the Centre International de Récherche Medicale de Franceville, Gabon, West Africa, December 1979. Pp.173. ISBN 0-906545-04-8. (The Biochemical Society, London: 1980.) £19, $45.

SUZUKI, D.T., GRIFFITHS, A.J.F. and LEWONTIN, R.C. An Introduction to Genetic Analysis. 2nd Edn. Pp.911, plus Solutions Manual and Instructor's Guide. ISBN 0-7167-1263-6. (W.H. Freeman: 1981.) $22.95.

SWAIN, T. and KLEIMAN, R. (eds). The Resource Potential in Phytochemistry. Recent Advnces in Phytochemistry, Vol.14. Proceedings of the 1979 Annual Meeting of the Phytochemical Society of North America, held at the Northern Illinois University, DeKalb, Illinois, August 1979. Pp.215. ISBN 0-306-40572-5. (Plenum: 1980.) Np.

TAYLOR, T.N. Paleobotany. An Introduction to Fossil Plant Biology. Pp.589. ISBN 0-07-062954-4. (McGraw-Hill: 1981.) $29.95.

THEODOR, O. Diptera: Asilidae. Fauna Palaestina Insecta II. Pp.448. ISBN 965-208-015-2. (The Israel Academy of Sciences and Humanities, Jerusalem: 1980.) $32.

USDIN, E., KVETNAŇSKÝ, R. and KOPIN, I.J. (eds). Catecholamines and Stress: Recent Advances. Developments in Neuroscience, Vol.8. Proceedings of the 2nd International Symnposium held in Smolenice Castle, Czechoslavikia, September 1979. Pp.618. ISBN 0-444-00402-5. (Elsevier North-Holland: 1980.) $65, $72.25. outside U.S.A.

VILCEK, J., GRESSER, I. and MERIGAN, T.C. (eds). Regulatory Functions of Interferons. Annals of the New York Academy of Sciences, Vol.350. Pp.641. Hbk ISBN 0-89766-089-7; pbk ISBN 0-89766-090-0. (The New York Academy of Sciences, New York: 1980.) Hbk $124; pbk Np.

WATSON, L. and DALLWITZ, M.J. Australian Grass Genera: Anatomy, Morphology, and Keys. Pp.209. Pbk ISBN 0-908160-95-X. (The Australian National University Research School of Biological Sciences, Canberra: 1980.) Pbk $12.

WATTERSON, D.M. and VINCENZI, F.F. Calmodulin and Cell Functions. Annals of the New York Academy of Sciences, Vol.356. Hbk ISBN 0-89766-101-X; pbk ISBN 0-89766-120-8. (The New York Academy of Sciences, New York: 1980.) Hbk $86, pbk Np.

WEIR, D. and SCHAPIRO, M. Circle of Poison. Pesticides and People in a Hungry World. Pp.101. Pbk ISBN 0-935028-09-9. (Institute for Food and Development Policy, San Francisco: 1981.) Pbk $3.95.

WHITTALL, J.D. People and Animals. Pp.154. ISBN 0-905225-03-1. (The National Anti-Vivisection Society, London: 1981.) Hbk £4.50; pbk £2.50.

WOODS, S.E. Jr. The Squirrels of Canada. Pp.199. ISBN 0-660-10344-3. (The University of Chicago Press: 1981.) £18.

WUNDERLICK, K. and GLOEDE, W. Nature as Constructor. Pp.196. ISBN 0-668-05102-7. (Arco Publishing, New York: 1981.) Np.

ZIMEN, E. The Wolf. His Place in the Natural World. Pp.373. ISBN 0-285-624-113. (Souvenir, London: 1981.) £9.95.

Applied Biological Sciences

BORSELLINO, A. *et al.* (eds). Developments in Biophysical Research. Proceedings of a Congress sponsored by the Italian Society for Pure and Applied Biophysics, held in October, 1979 in Parma, Italy. Pp.365. ISBN 0-306-40627-6. (Plenum: 1980). Np.

EVERHART, W.H. and YOUNGS. W.D. Principles of Fishery Science. 2nd Edn. Pp.349. ISBN 0-8014-1334-6. (Cornell University Press: 1981.) £10, $20.60.

GLOVER, S.W. and HOPWOOD, D.A. (eds). Genetics as a Tool in Microbiology. 31st Symposium of the Society for General Microbiology held at the University of Cambridge, April 1981. Pp.427. ISBN 0-521-23748-3. (Cambridge University Press: 1981.) £27.50.

KLEIN, G. and WEINHOUSE, S. (eds). Advances in Cancer Research. Vol.33. Pp.325. ISBN 0-12-006633. (Academic: 1980.) $37.50.

LUNDHOLM, B. and STACKERUD, M. (eds). Environmental Protection and Biological Forms of Control of Pest Organisms. Ecological Bulletins, No.31. Pp.171. ISBN 91-546-0280-7. (Swedish Natural Science Research Council, Stockholm: 1980.) 110 Skr, $25.

MILUNSKY, A., FRIEDMAN, E.A. and GLUCK, L. (eds). Advances in Perinatal Medicine, Vol.1. Pp.438. ISBN 0-306-40482-6. (Plenum: 1981.) $35.

NAKAMURA, R.M., DITO, W.R. and TUCKER, E.S. (eds). Immunoassays: Clinical Laboratory Techniques for the 1980's. Laboratory and Research Methods in Biology and Medicine, Vol.4. Proceedings of the 2nd Annual Conference, San Diego, California, February 1980. Pp.464. ISBN 0-8451-1653-3. (Alan R. Liss, New York: 1980.) $58, Dfl. 174.

TORSOLI, A., LUCCHELLI, P.E. and BRIMBLECOMBE, R.W. (eds). Further Experience with H_2-Receptor Antagonists in Peptic Ulcer Disease and Progress in Histamine Research. European Symposium Capri, October 1979. Pp.370. ISBN 90-219-9467-4. (Excerpta Medica, Amsterdam: 1980.) $63.50, Dfl. 130.

Psychology

BECHTEREVA, N.P. .(ed.). Psychophysiology Today and Tomorrow. Pp.263. ISBN 0-08-025930-8. (Pergamon: 1981.) £25, $60.

BELL, P.B. and STAINES, P.J. Reasoning and Argument in Psychology. Pp.217. Pbk ISBN 0-7100-1712-4. (Routledge and Kegan Paul: 1981.) Pbk £4.95.

BOTT, M. and BOWSKILL, D. The Do-It-Yourself Mind Book. Pp.205. ISBN 0-7045-03123. (Wildwood House, London: 1981.) £7.95.

BUGELSKI, B.R. and GRAZIANO, A.M. The Handbook of Practical Psychology. Pp.296. Hbk ISBN 0-13-380600-6; pbk ISBN 0-13-380592-1. (Prentice-Hall International: 1981.) Hbk £9.05; pbk £4.50.

DEREGOWSKI, J.B. Illusions, Patterns and Pictures: A Cross-Cultural Perspective. Pp.220. ISBN 0-12-210750-0. (Academic: 1980.) £14.80, $34.50.

EISLER, R.M. and FREDERIKSEN, L.W. Perfecting Social Skills. A Guide to Interpersonal Behaviour Development. Pp.226. ISBN 0-306-40592-X. (Plenum: 1981.) $18.95.

HASSIBI, M. and BREUER, H. Jr. Disordered Thinking and Communication in Children. Pp.207. ISBN 0-306-40490-7. (Plenum: 1980.) $25.

KLAUSMEIER, H.J. and SIPPLE, T.S. Learning and Teaching Concepts: A Strategy for Testing Applications of Theory. Pp.228. ISBN 0-12-411450-4. (Academic: 1980.) $22.50.

LEVY, N.B. (ed.). Psychonephrology, 1. Psychological Factors in Hemodialysis and Transplantation. Pp.287. ISBN 0-306-40586-5. (Plenum: 1981.) $25.

McCORMACK, M.K. Prevention of Mental Retardation and other Developmental Disabilities. Pediatric Habilitation Series, Vol.1. Pp.680. ISBN 0-8247-6950-3. (Marcel Dekker: 1980.) SwFr. 110.

NEHER, A. The Psychology of Transcendence. Pp.361. Hbk ISBN 0-13-736652-3; pbk ISBN 0-13-736645-0. (Prentice-Hall: 1980.) Hbk np; pbk np.

PIAGET, J. Intelligence and Affectivity: Their Relationship During Child Development. Pp.75. ISBN 0-8243-2901-5. (Annual Reviews, California: 1981.) $8, $9 (elsewhere).

RACHLIN, H. Behaviorism in Everyday Life. Pp.186. Hbk ISBN 0-13-074583-9; pbk ISBN 0-13-074575-8. (Prentice-Hall International: 1981.) Hbk £8.40; pbk £3.20.

REESE, H.W. and LIPSITT, L.P. Advances in Child Development and Behaviour, Vol.15. Pp.263. ISBN 0-12-009715-X. (Academic: 1980.) $28.

SOKAL, M.M. (ed.). An Education in Psychology. James McKeen Cattell's Journal and Letters from Germany and England, 1880-1888. Pp.372. ISBN 0-262-19185-7. (The MIT Press: 1981.) $30.

SPRINGER, S.P. and DEUTSCH, G. Left Brain, Right Brain. Pp.243. Hbk ISBN 0-7167-1269-5; pbk ISBN 0-7167-1270-9. (W.H. Freeman: 1981.) $15.95, $7.95.

WOLMAN, B.B. Contemporary Theories and Systems in Psychology. 2nd Edn. Pp.639. ISBN 0-306-40515-6. (Plenum: 1981.) $42.50.

YENI-KOMSHIAN, G.H., KAVANAGH, J.F. and FERGUSON, C.A. (eds). Child Phonology. Vol.1, Production. Pp.304. ISBN 0-12-770601-1. (Academic: 1980.) $27.

YOUNG, H.B. and FERGUSON, L.R. Puberty to Manhood in Italy and America. Pp.283. ISBN 0-12-773150-4. (Academic: 1981.) $26.

Sociology

COELHO, G.V. and AHMED, P.I. (eds). Uprooting and Development: Dilemmas of Coping with Modernisation. Pp.538. ISBN 0-306-40509-1. (Plenum: 1980.) $27.50.

LYTTON, H. Parent–Child Interaction: The Socialisation Process Observed in Twin and Singleton Families. Pp.364. ISBN 0-306-40521-0. (Plenum: 1980.) $35.

OHUCHE, R.O. and OTAALA, B. (eds). The African Child and His Environment. UNEP Studies, Vol.3. Pp.97. ISBN 0-08-025671-6. (Pergamon: 1981.) £6.25, $15.

RIBBENS, G.F. Pattern of Behaviour. A Comparative Sociology Text for the Critical Student. Pp.296. Flexi ISBN 0-7131-0412-0. (Edward Arnold: 1980.) Np.

SALZINGER, S., ANTROBUS, J. and GLICK, J. (eds). The Ecosystem of the "Sick" Child. Implications for Classification and Intervention for Disturbed and Mentally Retarded Children. Pp.308. ISBN 0-12-617250-1. (Academic: 1980.) $25.

SCUDDER, T. and COLSON, E. Secondary Education and the Formation of an Elite: The Impact of Education on Gwembe District, Zambia. Pp.190. ISBN 0-12-634280-6. (Academic: 1980.) $24.50.

SHANNON, T.A. (ed.). Bioethics: Basic Writings on the Key Ethical Questions That Surround the Major, Modern Biological Possibilities and Problems. Revised. Pp.632. Pbk ISBN 0-8091-1970-6. (Paulist Press, Ramsey, New Jersey: 1981.) Pbk Np.

VAZQUEZ, P.R. and MARGAIN, A.L. (eds). Industrial Design and Human Development. Proceedings of the XIth Congress and Assembly of the International Council of Societies of Industrial Design, Mexico City, October 1979. Pp.354. ISBN 90-219-0456-X. (Excerpta Medica, Amsterdam: 1980.) $73.25, Dfl.150.

General

SHANNON, T. A. and DIGIACOMO, J. D. An Introduction to Bioethics. Pp.154. Flexi ISBN 0-8091-2257-X. (Paulist Press, 545 Island Road, Ramsey, New Jersey 07446: 1980.) $4.95.

SILVEY, S. D. Optimal Design: An Introduction to The Theory for Parameter Estimation. Pp.86. ISBN 0-412-22910-2. (Chapman and Hall: 1980.) £6.50.

SIMPSON, G. G. Why and How: Some Problems and Methods in Historical Biology. Pp.263. Flexi ISBN 0-08-025784-4; hbk ISBN 0-08-025785-2. (Pergamon: 1980.) Flexi £9; hbk £18.

SLOAN COMMISSION ON GOVERNMENT AND HIGHER EDUCATION. A Program for Renewed Partnership. The Report of the Sloan Commission on Government and Higher Education. Pp.309. ISBN 0-88410-193-2. (Ballinger: 1980.) Np.

SORENSON, H. W. Parameter Estimation: Principles and Problems. Pp.381. ISBN 0-8247-6987-2. (Dekker: 1980.) SwFr. 95.

SPIEGLER, K. S. and LAIRD, A. D. K. (eds). Principles of Desalination. 2nd Edn. Part A. Pp.357. ISBN 0-12-656701-8. (Academic: 1980.) $41.

SPIEGLER, K. S. and LAIRD, A. D. K. (eds). Principles of Desalination. 2nd Edn. Part B. Pp.821. ISBN 0-12-656702-6. (Academic: 1980.) $52.

TAYLOR, R. and PRITCHARD, C. The Protest Makers. The British Nuclear Disarmament of 1958-1965, Twenty Years On. Pp.190. ISBN 0-08-025211-7. (Pergamon: 1980.) £10, $22.50.

TENNANT, A. The Guadalupe Mountains of Texas. The Elma Dill Russell Spencer Foundation Series, No. 10. Pp.168. ISBN 0-292-72720-8. (University of Texas Press: 1980.) £17.95.

THUILLIER, P. Le Petit Savant Illustré. Pp.118. Flexi ISBN 2-02-005699-2. (Seuil: 1980.) Np.

UNITED NATIONS ENVIRONMENT PROGRAMME. Choosing the Options: Alternative Lifestyles and Development Patterns. Pp.76. Flexi ISBN 92-807-1024-8. (United Nations Environment Programme, Kenya: 1980.) Np.

VAN FRAASSEN, B. C. The Scientific Image. Pp.235. Hbk ISBN 0-19-824424-X; pbk ISBN 0-19-824427-4. (Clarendon Press: 1980.) Hbk Np.; pbk £6.50.

VAVRA, R. The Love of Tiger Flower (around paintings by Fleur Cowles). Pp.46. ISBN 0-00-216208-3. (Collins: 1980.) £5.50.

WANG, N. T. (ed.). Business With China: An International Reassessment. Pp.145. ISBN 0-08-026292. (Pergamon: 1980.) $25.

WEST, J. M. Basic Corrosion and Oxidation. Pp.247. ISBN 0-85312-196-6. (Wiley: 1980.) £19.50.

WISE, K. A. J. (ed.). Records of the Auckland Institute and Museum, Vol. 16 for 1979. Pp.205. Flexi ISBN 0067-0464. (Auckland Institute and Museum, Private Bag, Auckland, 1, New Zealand: 1980.) Np.

WOOLF, H. (ed.). Some Strangeness in the Proportion: A Centennial Symposium to Celebrate the Achievements of Albert Einstein. Pp.539. ISBN 0-201-09924-1. (Addison-Wesley: 1980.) $43.50.

nature

Pure Chemistry Library
SCIENCE COLLEGE
[illegible]
[illegible]

18 June 1981

Making Israel atone for Tamuz

The chief casualty of the Israeli raid on the Iraqi reactor complex at Tamuz is the international reputation of the Non-Proliferation Treaty. Iraq, a ratified signatory of the treaty since 1969, in 1977 agreed with the International Atomic Energy Agency that all its nuclear operations should be open to formal inspection. What Israel has now done is to demonstrate to Iraq, and to other Arab and Muslim states, that formal obligations such as these provide no protection from Israeli interdiction. Even the government of Egypt, for the time being in Mr Menachem Begin's good books, must be wondering whether it will be immune from a similar attack when the power reactor to be built under an agreement with the United States, again under the aegis of the Non-Proliferation Treaty, is nearing completion. If the present Israeli mood survives the coming general election, the agreements carefully negotiated between Israel and Egypt could be put on the back burner for the sake of a pre-emptive strike against Egyptian competence in civil nuclear technology. Before that happens, and for the sake of repairing the political damage already done in the Middle East, the interested parties, but especially the governments of France and the United States, should look for a remedy that will strengthen the Non-Proliferation Treaty. Is it too much to ask that Israel itself should now be made to sign the treaty?

Israeli policy on nuclear technology in the past twenty years is not at first sight encouraging, yet the events of the past two weeks are an opportunity as well as a disaster. Ever since the construction of the Israeli reactor at Dimona in the Negev in the early 1960s, Israel has enjoyed the benefits of being a putative nuclear power. Israel's neighbours have been encouraged to suppose that Israel has a modest stock of nuclear weapons that could and would be used in anger in an emergency. This reputation has served Israel well. Some neighbours of Israel, Egypt and Syria in particular, have no doubt had to take into account the possibility of nuclear retaliation when there has been trouble in the Sinai or on the Golan Heights. On the other hand, Israel has escaped the contumely that would follow a tangible demonstration that it has nuclear weapons at its disposal. As things are, nobody knows for sure whether Israel has nuclear weapons or not — the only compensation is that if there are Israeli bombs, Israeli generals must be less confident than their opposite numbers elsewhere that they will function as intended. What are the chances that in the aftermath of the Tamuz raid, it will be possible to force Israel to come clean?

The Israeli justification of the raid on Iraq is paper-thin. That the raid was an open breach of international law is not disputed. That would have been the case even if, as the government of Israel claims, there were evidence that Iraq had been planning to use the 70 megawatt reactor being built under its agreement with France as a source of military plutonium. But in reality, whatever dreams the reactor managers of Tamuz may have been cherishing, the simple truth is that Iraq could not have made weapons with plutonium from its new reactor without the International Atomic Energy Agency knowing well in advance. The agreement between Iraq and France requires that enriched uranium for the reactor should be returned to France for reprocessing, while an attempt to generate plutonium from natural uranium (supposedly bought from Portugal) would surely have been spotted by the least competent of the inspectors from Vienna. Only in two respects could the developments in Iraq have been a military threat to Israel. Iraq might at some point in the future have abrogated its adherence to the Non-Proliferation Treaty, and set about using the reactor as a source of military plutonium. So far, no state has taken such a step, and while the major powers remain united behind the treaty none is likely to dare. The result of abrogation would be intolerable international pressure, the risk of not surviving the period between abrogation and becoming a nuclear power — and most probably a more legitimate version of last week's Israeli raid. Alternatively, Iraqi expertise might have been used as a way of helping like-thinking states, not signatories of the Non-Proliferation Treaty, to make nuclear weapons. Pakistan is most often mentioned. Strictly, assistance of this kind is against the rules. In practice, violations are harder to detect or at least to prove. But again, there is no reason why the government of Israel should have taken it on itself to pre-empt such developments.

But why should Iraq be interested in nuclear power? Should not oil-rich states content themselves with drilling holes in the ground? That is a second line of defence of the Israeli raid advanced by Mr Begin and his apologists elsewhere. The simple answer is "Why not?" Iraq is not, in any case, the first oil-rich state to take an interest in civil nuclear power. Iran in the Shah's day was an apparently enthusiastic buyer of nuclear reactors from France and the United States. The logic was simple. Why use chemically sophisticated hydrocarbons as a source of heat when they could be used as starting points for chemical manufacturing instead? More crudely, why not use what seemed once to be cheap nuclear energy to generate electricity, allowing large amounts of oil to be sold at $40 a barrel or whatever it will make on a seller's market? There is no reason why Iraq should not (at least in 1977) have nursed such an ambition. It is in any case as plausible that Iraq may have been calculating — that the calculation may have been wrong is irrelevant — that a dash of civil nuclear technology might be good for a technologically backward state. That prescription has worked to some incalculable effect in India. And if it is true that Iraq and France have been less open about their agreement on the Tamuz reactors (there is — or was — also a 600 kilowatt reactor) than they should have been, who, in Israel at least, can complain if the parties to the agreement kept it to themselves for fear of provoking trouble from Israel?

For the time being, responsibility for responding to the Israeli raid rests with the governments of Iraq, France and the United States, in that order of immediacy. What should they do? Iraq, laughably enough, probably has a civil suit against Israel; in spite of the cumbersomeness of international legal processes, that should be pursued. There is no reason why the reactor at Tamuz should not in due course be replaced, and at somebody else's expense. For the new government of France, the raid on Tamuz could hardly have been more ineptly timed. The new president had plainly been looking forward to reconstructing French relationships in the Middle East, perhaps resurrecting some of the friendliness between France and Israel that ended disastrously at Suez in 1957; he will now have been given pause. There is, however, no reason to suppose that the new French government will be as enamoured as was M. Jacques Chirac, the Prime Minister of France in 1977, of assisting with the creation of independent nuclear enterprises elsewhere. The chances are that President Mitterrand will confine himself, for the time being, to the moderate protests of the past few days. Further ahead, however, there is at least a chance that the new France may be tempted to sign the Non-Proliferation Treaty. M. Mitterrand has cautiously avoided all but passing references to the subject. Somebody should ask him what is in his mind.

The principal burden will, however, rest with the United States, hamstrung as always by what is called the Zionist vote. President

0028-0836/81/250523-02$01.00

Reagan has found a neat way of shifting the immediate burden of awkward choice from the White House to the Senate Foreign Relations Committee, which will have to confirm or deny his suspension of aircraft sales to Israel. Sooner or later, however, the buck will come to rest again in the Oval Office. At some point, the Administration will have to devise a new anti-proliferation policy — the Carter version is too inflexible, and ideologically unsuited to this Administration. With the Tamuz raid fresh in people's minds, Mr Reagan will be repeatedly reminded (by Arab states if not by his voters) that the aircraft which carried out the raid on Tamuz were supplied by the United States and would have suited well for the delivery of nuclear weapons. Does it make sense, people will be asking, that the United States should resist the proliferation of nuclear weapons but be prepared to supply delivery systems to putative nuclear powers such as Israel? And how can the United States, which has invested a vast amount of political capital in the past few years in the good cause of persuading other states to sign the Non-Proliferation Treaty, now give Israel (which has not signed) a licence to undercut that policy? It is not as if Israel's supporters, in the United States and elsewhere, seriously consider that Israel's safety depends on the undefined threat of retaliation with nuclear weapons. The supporters of Israel's continued independence would indeed be more robust — and more numerous — if there were not the continually nagging fear that help would be used eccentrically, even illegally. This is why the time has come for Israel to sign the Non-Proliferation Treaty — and for the United States to ensure that it is done.

Cheque-book academics in trouble

Stanford University seems ready to take a cautious step towards an accommodation between the principles of the academic life and the wish of some academics to turn their expertise to profit. The proposals (see page 526), on which the Stanford faculty will be asked to decide in a few months' time, have been made necessary by the apparently ceaseless spawning of commercial companies aiming to exploit genetic manipulation for commercial purposes and which include academic researchers as stockholders or even as entrepreneurs. The objective is to avoid as far as possible the conflicts of interest that arise when academics individually, or a university corporately, become dependent on commercial enterprises. The method is simple enough, almost deceptively so. The university will not (if the proposals are accepted) take an equity interest in commercial enterprises in which members of the faculty have a similar interest, while individual teachers will be required to disclose their outside interest whenever asked to do so by some more senior person, a dean or even the president of the university.

The university's proposed act of self-denial is virtually a replica of the decision reached at Harvard last October, when the faculty there persuaded the university administration not to go ahead with a proposal to accept (without payment) a block of shares in a genetic engineering company being organized by Professor Mark Ptashne. The arguments then, as now at Stanford, were straightforward. With a substantial minority stake in a successful business, a university may find it hard to avoid giving those who are nurturing the golden goose unfair advantages (compared with colleagues) in matters such as appointments or even the allocation of space. This, however, is the least of the worries likely to confront universities with equity stakes in companies. What, for example, would happen if such a company were in financial trouble? Would it be possible for a university to resist pleas for just a little help, a little cash perhaps, or forbearance with bills inconveniently receivable? Harvard and now Stanford are right to decide that the risks are too great too be stomached. Neither university has so far taken a view on the propriety of the role of academics as spare-time entrepreneurs.

The obvious reason is that most university faculties would be hopelessly divided on any such proposition. So much is clear from the way in which Stanford has sidled up to the proposition — not yet agreed — that members of the faculty should disclose their outside interests to the academic administration whenever it seems possible that a conflict of interest has arisen. Why wait until then? Why not simply agree that academics have a duty to disclose all their business interests to their immediate colleagues, not merely to august administrators with an ear for gossip? The obvious drawback in the arrangement now proposed at Stanford is that academics with outside interests (not necessarily commercial) are not the best placed to decide that a conflict has arisen with their academic responsibilities. Colleagues are likely to be more sensitive. They should be given the right to express reasonable doubts, no doubt within a framework of understanding that the process of innovation is often furthered by means of links, consultancy agreements for example, between academics and industry.

Stanford, then, should opt for full disclosure, at the level of the department, of academics' outside interests. The outcome would probably be surprisingly dull. Most people, it would emerge, have modest consultancy arrangements with the odd company or two, consider themselves ill-rewarded for the advice they give and make up for that by skimping on what they do. Only occasionally would it emerge that an academic's commitment to some commercial company was offensive, perhaps because of being exclusive to one external organization, perhaps (and worse) because it endangered the proper balance of a research programme or the education of graduate students.

What, against this background, is to be made of the growing company of academic entrepreneurs, the people who help to start commercial companies, often by telling the tale of the promise hidden in their research to the moneybags who must in the end provide what is called the "venture capital"? The fashion, in the past few months, for taking on this role has been most conspicuous in the application of molecular biology but is unlikely to be thus restricted for very long. Now that molecular biologists have shown how it is possible to make paper fortunes when a go-go enterprise is launched on an unsuspecting stock market, it is unlikely that computer scientists, organic chemists and people in general will hang back. The simple truth, however, is that there is a difference between a consultant — a *per diem* person — and one who helps to start a company. The former can walk away from disaster with an easy conscience; the latter must redouble his efforts when disaster threatens. The conclusion must surely be that an academic who opts for the role of entrepreneur must at the same time opt out of tenure. In the last resort, the argument must go, the faculty cannot reasonably command such a person's loyalty.

In the United States, the seemly regulation of these questions would be simpler if it had not in the past few years been complicated by the amendments of the Patents and Trademarks Act. In the old days, the outcome of a federally-sponsored research programme was in the public domain and thus up for grabs from anybody with the wit to recognize its importance. Now, universities at which research projects are undertaken have a right, even a responsibility, to patent what they can and to exploit it commercially. The result is often a nonsense, tacitly acknowledged as such in the burgeoning field of monoclonal antibodies (see page 525). The snag is that universities by their constitution are not much good at playing patent agents, but (given the continuing threat of shrinking funds) awkwardly eager to do so. The message seems to have spread quickly to teaching staffs. All concerned (the federal agencies in the United States included) should brood on what they have done before things go too much further. Elsewhere, in Britain for example, academics will soon be envying what seems to be the luxury of making money decisions. But worrying about earning money can be as sapping of scholarship as knowing that there is none.

0028-0836/81/250524-01$01.00

Hoechst makes deal with Mass. General

House panel examines overseas ties

Washington

University involvement with the commercial exploitation of genetic engineering is well on the way to becoming an issue between the United States government and the universities. Last week, at a meeting of Congressman Albert Gore's oversight subcommittee of the House Committee on Science and Technology, representatives of the Massachusetts General Hospital were hauled over the coals because of the hospital's agreement with Hoechst AG under the terms of which the German chemical company will provide $50 million over the next ten years to run a molecular biology laboratory at the hospital, and will have first refusal of any patent licences that may result.

At the hearings, Dr Ronald Lamont Havers, Director of research at the hospital, demurred when asked to provide the subcommittee with a copy of the agreement it has signed with Hoechst. The subcommittee's chairman left the hospital and everybody else within earshot in no doubt of his belief that universities sustained for the past twenty years by federal grants should be careful not to give overseas enterprises favoured access to their expertise. There is now some talk that the subcommittee will subpoena the agreement between the Massachusetts General Hospital and Hoechst if the document does not turn up in the mail before too long.

The grant from Hoechst will be used to found a department of molecular biology at the hospital. Dr Howard M. Goodman of the University of California, San Francisco, will be the director of the enterprise, the staff of which is expected to grow to about 50 by 1983 and thereafter to 100. The hospital insists that the programme of research will be determined by the hospital alone, while the work of the department will be reviewed by a committee of six scientists independent of Hoechst. The operation of the agreement between the company and the hospital will be kept under review by a joint committee of three senior managers of the company and three trustees from the hospital.

The declared objective of the new department is to apply the techniques of molecular biology to the treatment of disease. The hospital says that all appointments in the new department will be made according to established academic procedures, that individual scientists will be free to publish how and when they choose provided that the hospital authorities are informed in advance, and that there will be no restrictions on collaboration with scientsts elsewhere.

According to the hospital, patentable discoveries will be patented by the hospital, and the benefits shared between the hospital and the inventor on the basis of existing rules. Hoechst will have a right to a licence to any patent springing from research which it has sponsored, while the royalty rate negotiated "will reflect the financial contribution of the company".

Two potentially contentious aspects of the agreement are that the company will be free to decide whether or not particular research projects should be funded out of the $50 million set aside, and that members of the new department working on Hoechst-sponsored projects will not be free to consult with other companies. The agreement is automatically renewable if not terminated after ten years.

It has also been agreed that there will be a public seminar once a year to which some Hoechst scientists will be invited, and that the company will have a right to send up to four of its people to the hospital for training at any time.

Congressman Gore is unlikely to let these issues fade away. After last week's hearings, he said that "the questions that

NIH plan new overhead calculation

Academic research administrators will soon have to learn new algorithms for calculating how much overhead to charge against successful applications for research grants and contracts with the National Institutes of Health (NIH). So much is clear from the meeting last week (8–9 June) of the Institute's Advisory Committee, at which the task force led by Professor Samuel O. Thier of Yale University that has been brooding on the question of "institutional support" (bureaucratic jargon for overhead) promised to produce at the next meeting in October several proposals for reorganizing the present system.

Objections to the present system are several. Research grant proposals approved by study sections of NIH (the "peer review" committees) are made more costly by an amount negotiated between NIH and the recipient institutions intended to cover the cost of supporting the research concerned. On the average, NIH overheads cost 30 per cent of total expenditure on research grants and contracts. There are wide geographical variations, with the universities in the north-west of the United States successfully claiming larger percentages for institutional support than universities elsewhere.

The stimulus for the present review has come from federal agencies and universities, both equally appalled at the difficulties of carrying out the detailed accounting for research grant expenditure now required of them. But NIH also have an interest in heading off trouble about institutional support from the Reagan Administration, which has already shown its hand by removing from the NIH budget institutional support accompanying awards of postdoctoral fellowships.

The Thier task force plans to spend the summer assessing the merits of alternative mechanisms for providing institutional support. The favoured alternative to the present system appears to be the "fixed obligation" formula, under which research grants and contracts would be awarded on what is essentially a fixed-price basis, and on which federal investigation would be limited to a simple verification that direct and indirect costs have been properly incurred.

At the next meeting of the Advisory Committee in October, the task force will suggest that experiments should be carried out with several of the alternative methods of financial research projects. It is, however, unlikely that all the loose ends can be tied up by October — it is not, for example, at this stage clear whether the overhead element in fixed obligation grants should be assessed by a peer review committee and, if so, whether should be the appropriate study section.

Last week's meeting of the NIH Advisory Committee also wrestled a little inconclusively with the issue of patent rights in biological innovations.

One curious development to come to light was that many scientists are cheerfully ignoring the provisions of the Patents and Trademark (Amendments) Act of 1979. This legislation gives universities at which research grants are held the right to patent and exploit new developments, reserving to the federal government a non-exclusive right to a licence. Under the amended law, those holding grants from NIH are required to report all patentable developments.

These provisions have already been found unworkable in the development of monoclonal antibodies, each one of which may be potentially patentable. Given that a productive laboratory may expect to develop several hundreds of monoclonal antibodies a year, and that the cost of patent protection is a minimum of $2,000, it seems to have been tacitly assumed that only those monoclonal antibodies likely to be commercially important deserve the investment of $2,000-plus.

0028-0836/81/250525-02$01.00

will confront our civilization as a result of genetic engineering" will be the most difficult ever, and that their solution will only be complicated if those responsible for original discoveries make commercial links with industry.

The subcommittee last week found a more congenial witness in Dr Donald Kennedy, president of Stanford University, who earlier this year was advocating a "second Asilomar conference" on the problems of university-industry relationships. Kennedy, who more recently has been back-pedalling on this proposal, told the subcommittee last week of the ambivalence of research universities towards proposals for industrial support in the present climate of stagnant federal budgets. Well-ordered agreements, he said, could be beneficial, but he also pointed to the dangers of conflicts of interest, secrecy and the distortion of research training.

Universities are perhaps more than usually aware of these problems because of the advocacy of the virtues of industrial support for the research universities by the National Science Board.

Yugoslavs to resign

Dr Pavle Savic, President of the Serbian Academy of Arts and Sciences since 1971, has resigned in protest against the conduct of recent elections to the Academy. His resignation, which takes effect from the academy meeting of 28 May, came as no surprise — it had been openly discussed in the Yugoslav press for the preceding three weeks.

According to Dr Savic, he resigned because of the election of someone who had served several years' rigorous imprisonment as a "cominformist". He had been informed of this one month before the elections. On checking the allegation and finding it true, Dr Savic notified the appropriate Academy officials. But the election went ahead and the "cominformist" was elected. The next day, Dr Savic announced his intention to resign. Two vice-presidents and the general secretary of the academy also resigned; they, however, will serve out their full term of office (until the end of the year) to maintain continuity of the academy's activities.

Dr Savic's walk-out, seven months before his term of office ended, is widely seen as a protest, not merely against the election, but against the prevailing atmosphere in the academy. A physicist, he had hoped, to build up the research side of the academy, concentrating on modern trends including laser physics, solid-state electronics and molecular biology. He was particularly keen to bring in younger scientists into the academy's institute. Most members, however, showed little enthusiasm for his plans.

Vera Rich

Exploiting research

Stanford rules

Washington

A proposal for regulating the commercial activities of members of the Stanford University faculty is to be put to the Academic Senate at its meeting in September. One of the proposals is that members of the faculty should disclose details of consultancy agreements with commercial organizations if there are reasons to expect that conflicts of interest have arisen.

The new scheme has been worked out by a committee under Dr Ingram Olkin, professor of statistics and education at Stanford, which has recommended that the university should not invest in commercial ventures set up to exploit the results of research carried out at Stanford "if any current Stanford faculty member participates" either as a shareholder or as a manager. Similarly, the university would not accept stock in such a company as part of a licensing agreement.

The proposals have been stimulated by the recent interest in the commercial exploitation of genetic manipulation, to which several members of the Stanford faculty are in a position to contribute. But it is acknowledged that similar problems may arise in other fields, the development of computer software for example.

The case for a distant relationship between the university and commercial enterprises in which faculty members are involved is based on the fear of the conflicts of interest that could arise in making faculty appointments and promotions, allocating space and settling faculty salaries, the danger that graduate education might be distorted and the possibility that the university's external reputation might be affected adversely.

The committee also advocates a change in the method by which the rewards from agreements for the licensing of patents are shared out. At present, if Stanford research leads to a patent, the income is divided equally between the researcher, his or her department and the university as a whole. It is now suggested that if the income from some invention should exceed one half of a department's annual budget, the two-thirds of the royalty income not due to the inventor should be paid into a research fund from which all departments at Stanford could hope to benefit.

The committee has obviously been cautious, perhaps even overcautious, in its approach to consultancy agreements. At present, academics are required to disclose the amount of time they spend in this way, which must not exceed 13 days a quarter. There is no mechanism for monitoring this gentleman's understanding. In future, if the faculty agrees, faculty members may have to disclose who they work for, and for how long.

Tropical disease finance

World Bank acts

After a plea by Mr Robert McNamara, who resigns as president of the World Bank on 30 June, the bank governors are expected to agree this week to contribute $2.48 million (£1.25 million) to the Special Programme for Research and Training in Tropical Diseases (TDR) of the World Health Organization (WHO).

Previously, the bank has been willing to act merely as a financial adviser and banker to TDR, as one of the programme's three co-sponsors. (The others are WHO and the United Nations Development Programme.) The bank's main concern is with development loans, and it has been exploring for more than a year how to refine its policy towards the direct support of research. That review is not complete, but McNamara thought it well enough advanced that a commitment should now be made to the tropical disease programme. He carried his executive board with him, and their recommendation has now to be approved by the governors — the nations which fund the bank. There is not expected to be any opposition.

The support for TDR comes at an opportune time. First firmly established in 1977, and dealing with six diseases — malaria, schistosomiasis (bilharzia), trypanosomiasis (sleeping sickness), leishmaniasis, leprosy and filariasis — TDR's funds first grew rapidly but have now become roughly static at about $24 million (£12 million) a year. Moreover Britain, which was contributing around £600,000 a year (but actually received back more in research grants, according to WHO), announced last December that it could not pay its 1981 contribution — shocking the 25 other contributing governments, many of which see little or no return in terms of research carried out in their countries.

According to the Canadian chairman of the TDR Scientific and Technical Advisory Committee, Dr A. B. Morrision, the effectiveness of the programme could be "crippled in the short term" if inflationary pressures are not compensated for; and in the long term, additional resources must be mobilized to permit large-scale trials of new tools for prevention and treatment now being developed.

Even so, Britain's contribution to TDR for the present year is definitely cancelled, says the Overseas Development Administration (ODA). And although next year's is "under review", the ODA's £1,000 million of aid money this year will fall by 5 per cent a year for the next three years, making a recovery of the TDR contribution less and less likely. This should not be taken as an implicit criticism of TDR, say ODA officials. On the contrary, the programme is the best managed of all the WHO special programmes, they say, thus putting a cloud over the two remaining programmes still

0028-0836/81/250526-02$01.00

supported by ODA.

British policy is to cut multilateral, special programme aid hardest, for the government's aim is to maintain direct core subscriptions to international agencies (such as the World Bank and the United Nations Development Programme), while stressing bilateral aid which is inevitably under closer British political and economic control.

Robert Walgate

Armadillos fight leprosy

The safety and ethics panel of the World Health Organization (WHO) is likely to agree within the next few months on immunization trials with a new leprosy vaccine made from the liver and spleen of infected armadillos. The trials would be conducted from January to September 1982 in the United States, United Kingdom and Scandinavia to avoid criticisms of uncontrolled drug testing in third world countries.

The objective will be to induce 85–90 per cent immunity to the causative agent of leprosy (*Mycobacterium leprae*) as judged by immune reactions to lepromin, a preparation of killed *M. leprae* obtained from infected human ear-lobes.

The vaccine has been a long-standing objective of the WHO Special Programme for Research and Training in Tropical Diseases (TDR). The snag is that it depends on culturing the bug in armadillos, whose low body temperature allows growth of the organisms, and some have questioned whether it will ever be possible to breed enough armadillos to vaccinate the 2,000 million or so of the human population potentially at risk.

Dr Barry Blum of the Albert Einstein College of Medicine in New York, who heads the TDR leprosy immunization project, says, however, that armadillos are such prolific generators of *M. leprae* that the numbers should be sufficient. A mixed vaccine made from human tissue yields immunity with a single dose of 6×10^8 killed bacteria. An armadillo three years after infection produces about 2.5×10^{12} bacteria in the liver and spleen. If the vaccine produced from them is as effective as the human one, each armadillo should yield 4,000 doses. WHO now has 250 animals used for research and the forthcoming trial, but there is a population of 10 million armadillos in the southern United States.

The adequacy of the supply depends directly on the dose required to produce immunity. One shot must be 85–90 per cent effective to avoid the need for repeated injections. The availability of the vaccine would also depend on the efficiency of extraction from armadillo tissue, a problem which has been given to the Wellcome Laboratories in the United Kingdom to resolve. Then leprosy is such a slow-moving disease that proof in the field against infective organisms will take ten years.

Robert Walgate

US solar energy

No silver lining

Washington

The Solar Energy Research Institute at Golden, Colorado, has been flung further into limbo by the arrival of the new Administration, which is now trying to respond to a demand from the General Accounting Office that the objectives of the institute should be spelled out. Most probably the institute, once the spearhead of President Carter's ambition that a fifth of United States energy consumption at the end of the century should be solar energy, will be asked to concentrate on long-term research, abandoning the commercial demonstration of solar energy devices. The present budget of $130 million is likely to be cut by a half, and between 100 and 200 of the present staff of 770 will probably lose their jobs.

The brief history of the institute, set up in 1977 and managed for the Department of Energy by the Mid-West Research Institute of Kansas, is a tale of muddle. Originally, there was to have been a staff of 1,200 engaged on long-term research, development and the commercial demonstration of devices able to win energy from the Sun. In part, the institute has been the victim of the success of some applications of solar energy. Solar water heaters, large and small, and small hydroelectric plants are multiplying throughout the United States — the Department of Energy is embarrassed by more than 1,000 applications from private companies and individuals to build hydroelectric plants. With the Administration committed to the view that commercial developments should be the responsibility of commercial organizations, not the government, the institute's demonstration programme has become vulnerable. So much seemed clear during a visit to Golden by officials of the Department of Energy last month.

Even when the axe falls, in the next few weeks, the rump of the institute is unlikely to be reassured about its future. Although the institute is now nearly four years old, the special building (on a 300-acre site offered by the Government of Colorado) seems as far off as ever. In a sharp letter to the Secretary of Energy at the end of April, the General Accounting Office asked that further design should be postponed until the department had decided what the long-term role of the institute should be, and that even the use of the 300-acre site for experimental rigs should wait on this definition. One snag is that the offer of the site expires in April 1982.

The letter is scathing about the planning of the building. Although authority for a laboratory was included in the 1974 Solar Energy Act, the estimated cost of $132 million (in 1978) was an unwelcome surprise. Successive Secretaries of Energy have sought to limit the cost of the buildings to $98 million (in 1979) and $75 million (in 1980). The General Accounting Office points out that the original scheme for making the building 80 per cent self-sufficient in its own energy requirements was based on the use of active and passive solar power devices which had not been proved and which were not all cost-effective. Meanwhile the institute's staff camps out in rented office accommodation costing $5 million a year, and has spent substantial sums on the conversion of offices to laboratories.

Since 1977, the institute's research programme has been trimmed by the removal of general responsibility for the exploitation of biomass, but it remains responsible for the use of methanol as a liquid fuel — a motor vehicle driven by hydrogen derived from the catalytic decomposition of methanol will be demonstrated later in the year. The institute is also confident that it will be able to produce silicon-based solar cells yielding power at a cost of 70 cents per watt by the end of this decade, roughly a tenth of the present cost of solar cells.

European Space Agency

Delays in space

For the second time this year, the European Space Agency (ESA) may be hamstrung by a change of plans within the United States National Aeronautics and Space Administration (NASA). This time, the possible casualty is Spacelab, ESA's manned space laboratory due to be launched on the space shuttle. Spacelab's first flight could be delayed, or its success reduced, if the Department of Defense persuades NASA to modify and thus delay the launch of the second of the two satellites intended to form the basis of a new data collection system.

The Tracking and Data Relay Satellite System (TDRSS) will eventually replace much of the worldwide network of tracking stations for relaying data from the space shuttle and satellites in low Earth orbit. Two tracking satellites, to be launched into geosynchronous orbits from early shuttle flights, were to transmit data from almost any position in a low Earth orbit to a central receiving station at White Sands, New Mexico. The original plan was to launch the first tracking satellite on the sixth shuttle flight, the second in June 1983 and Spacelab itself in September 1983. But now the Department of Defense has asked that extra coding should be built into the second satellite which could delay the launching until after that of Spacelab. One possibility is that the first Spacelab flight would be brought forward to June 1983 and the second tracking satellite launched only in September 1983.

If NASA accedes to this request, ESA would have the choice of either flying the first Spacelab payload with only one tracking satellite in service or of delaying the flight until the spring of 1984. One

0028-0836/81/250527-02$01.00

All aboard for Halley

The planned Soviet–French joint mission to Halley's comet in 1986 has been transferred, on the initiative of the Soviet Union, to the joint Comecon "Interkosmos" programme. At the same time, French participation will be largely reduced to the first stage of the mission, the fly-by of Venus.

The Halley probe is now to carry apparatus from Comecon countries, although their participation seems to be confined to the provision of equipment rather than the design of any individual experiments. Some countries may be assigned an individual responsibility for various sections of the mission — Hungary, for example, has been entrusted with developing the television system for monitoring the comet and transmitting the pictures to Earth.

The announcement of the Interkosmos trip to Halley's comet came at last month's meeting in Erevan of the Comecon permanent working group on space physics, only a few days after the successful landing of the Romanian cosmonaut, Dumitru Prunariu, brought to an end the programme of Comecon participation in missions to the Salyut space stations. Academician V.A. Kotel'nikov, chairman of Interkosmos, confirmed that no further joint manned flighs have been planned, although several unmanned joint missions are scheduled.

Vera Rich

tracking satellite would be able to cover the shuttle only in about half of its orbit and one third of Spacelab data could be lost. Extra recording equipment would be carried on board, but it is unlikely that the very high rate of data (100–150 megabits per second) from the microwave facility for Earth observations could be recorded. Spacelab users and ESA await NASA's final decision.

Meanwhile, NASA has been revising its schedule for shuttle flights up to 1985. Surprisingly, after the almost complete success of the first shuttle flight, delays are in store. Seven of the 44 flights scheduled before 1985 are to be postponed until the second half of the decade or even cancelled. Most other payloads are to be delayed. The reason for the delays, ranging from a few days to more than a year, are cuts in the Reagan budget, a more accurate estimate of the turn-round time and, most recently, delays in the manufacture of lightweight external fuel tanks. NASA says that delay in delivery of the tanks is largely responsible for the postponement of seven flights beyond 1985.

Three of the seven flights will effectively be cancelled. One will be saved by launching the two Galileo probes to Jupiter, more than 400 days later than originally planned, on a single shuttle flight. Another has been saved by the cancellation of the Venus-orbiting probe. And the third may be saved from the International Solar-Polar Mission if NASA fails to win approval for even a modified solar-polar spacecraft. (The ESA solar-polar spacecraft is included in the shuttle manifesto but with a launch date of May 1986, more than a year later than originally planned.)

Foreign and commercial users affected by the latest shuttle delays are being given until the end of this month to decide whether they would like to use Thor Delta launchers instead.

Judy Redfearn

European innovation

Brokers in demand

Luxembourg

The role of the European Commission in creating the right environment for innovation came up again at a symposium — sponsored by the Commission — in Luxembourg last week. The three-day conference was inspired by the resolution of EEC science ministers at the end of 1979 and follows on from a similar symposium on banking and innovation held last year.

Despite the ritual homage paid to the idea that small business will make good the job losses in large companies and industries, this year's symposium highlighted the gap between the researcher and the businessman. Participants repeatedly affirmed that basic and pure research in Europe too often fail to lead to commercial application.

The implication is that the dynamism of the small American business is not so easily copied in Europe. Not only is there a lack of venture capital but linguistic differences, the small size of national markets and technical barriers are further deterrents. In addition, the sheer mass and variety of the channels through which research results are made available bewilder rather than inspire industry. So this symposium edged round to the conclusion that what is required is a broker or intermediary who could select exploitable research and help to develop it commercially.

Some effective middlemen exist. In the United Kingdom, the National Research Development Corporation was set up in 1946. The Netherlands has the Eindhoven University of Technology, which has helped 200 small and medium-sized businesses in its first year of operation. In Austria the Innovationsgesellschaft is a venture bank specializing in helping inventors to launch new products. Belgium has several industrial research parks at the Catholic University of Louvain-La-Neuve and Vrije Universiteit Brussels.

The consensus of opinion at the symposium, however, was that these innovations are inadequate. One suggestion was for a European Technology Transfer Centre, another for a European Investment Bank specializing in research. Others proposed more bulletins on the lines of the French *Lettre des Sciences et Techniques* which selects and condenses reports of new discoveries. Many participants argued that the mentality of scientists in the field needs changing. They should publish information less for their peers than for the end user. Or they should themselves become entrepreneurs. The Commission will now consider how best to create the environment for the scientist/entrepreneur and for a European context for the scientific information broker.

Coincidentally, the European Commission's proposals for EEC programmes in the field of new technologies — microelectronics and telematics — are to be reexamined with greater interest, said Mr James Prior, the UK Secretary of State for Employment, after an unprecedented gathering of the EEC's finance and employment ministers in Luxembourg on 11 June with the aim of finding a solution to the growing unemployment problem. Sir Geoffrey Howe, the British Chancellor of the Exchequer, has also stated that the Commission is to bring forward new proposals to stimulate the expansion of small and medium-sized businesses.

If Mr Prior's words at the "Jumbo Council" are to be taken seriously, the Commission's work in the area of scientific communication is likely to be given a shot in the arm. The ministers agreed that there must be a greater coordination of the EEC's approach to new technologies.

Jasper Becker

Soviet scientists

Another trial

Dr Viktor Brailovskii, the Moscow cyberneticist faces trial this week on a charge of "disseminating fabrications . . . which defame the Soviet political and social system". Dr Brailovskii has refused a defence counsel, maintaining that nothing in his conduct over the past few years needs to be defended in court.

Dr Brailovskii was dismissed from his lectureship at the Moscow Radiotechnical Institute in 1972 when he applied, with his wife and son, to emigrate to Israel. In 1973, he and his wife Irina joined the "Sunday seminar on Collective Phenomena" organized by their friends, Dr Aleksandr Voronel', for "refusnik" scientists who wished to keep up some sort of intellectual life during the waiting period between applying for a visa (and subsequently losing their jobs) and actually being allowed to emigrate. From then on they were both frequently subjected to police harassment, which intensified when Dr Brailovskii became organizer of the seminar following the departure of Voronel' and his successor Dr Mark Azbel' to Israel.

In 1976, Dr Brailovskii was given permission to emigrate but his wife was refused on the grounds that she had access to secret information. (Her former immediate supervisor was prepared to deny this verbally, but not to commit himself in writing). On refusing to leave without his wife, Viktor Brailovskii had his permission to emigrate withdrawn in 1977.

In 1978, after a search of their apartment, his mathematical papers were confiscated and he was warned to stop lecturing on mathematics. In April 1980, the fourth in a series of conferences was to take place, basically an extension of the regular seminars to which foreign colleagues and sympathizers were invited. Before the 1980 conference Dr Brailovskii was arrested, questioned, and threatened with imprisonment if the meeting went ahead. Nevertheless it took place on schedule.

On 13 November 1980 Dr Brailovskii was arrested, and for several Sundays the police prevented anyone from going to his apartment. Dr Irina Brailovskaya, however, soon had the seminars going again, at first on Saturday evenings, and then, when the pressures lessened, at the original time on Sunday mornings. Dr Brailovskii, meanwhile, was held in Butyrki prison for more than 6 months (the official time for pre-trial detention is not more than three months, unless a special extension is sought).

Optical astronomy

Dutch sign on

Research councils in the Netherlands and the United Kingdom will sign an agreement today (18 June) which gives Netherlands astronomers 20 per cent of Britain's observing time at the Roque de los Muchachos Observatory (formerly called the Northern Hemisphere Observatory) in the Canary Islands.

In exchange for time on Britain's four planned telescopes — the 4.2-m William Herschel, 2.5-m Isaac Newton, a 1-m photometric telescope, and a 15-m millimetre-wave dish — the Dutch will provide 20 per cent of British costs and manpower in equipping and running the observatory.

The UK Science Research Council has estimated that the telescopes would cost £25 million in October 1979 prices, and that the running costs of the site will be of the order of £1 million a year.

The Netherlands has been hesitating over its contribution because of the sums involved, and also because it represents a certain shift of resources in favour of optical astronomy. The Dutch natural science research council (SWO) has previously done much of its optical astronomy through its membership of the European Southern Observatory, which has a fixed subscription; and Dutch radio-astronomers have voiced fears that the additional international commitment to the Roque will cut into their own budgets.

However, Dutch astronomers now agree that this level of commitment to Northern Hemisphere optical astronomy is necessary. The world-famous Westerbork radio telescope is, after all, in the Northern Hemisphere, and more direct links with optical observations would be useful.

The observatory they thus become part of was set up by international agreement between Spain, the United Kingdom, Sweden, and Denmark, and though the Dutch enter by "the back door" — the new agreement being between research councils and not nations — such deals were allowed for in the original treaty. **Robert Walgate**

Trademarks and patents

More authorities

Brussels

Members of the European Federation of Pharmaceutical Industries' Associations have expressed anger at the European Commission's proposals to create a Community trademark office. The drug industry, which accounts for 65 per cent of all trademarks registered, feels threatened by what it sees as an attempt to ride roughshod over its interests.

Last November the Commission brought out a directive aimed at harmonizing the member states' trademark laws, with proposals for the creation of a Community trademark and an EEC trademark office.

Trademarks are one way of ensuring protection against imitators of new medicines and guaranteeing that drugs are properly used. The pharmaceutical industry's federation, meeting in Brussels in May, concluded that placing the registration of trademarks in the hands of the Commission could lead only to confusion, as national authorities will operate concurrently with the Commission. The federation would prefer to wait until national rules are more in line with each other before setting up a European trademark office.

There is much confusion in Europe not only over trademarks but over patents and the protection of intellectual property in general. This is an economic as well as a legal problem, for representatives of industry continually stress that inadequate protection of new products puts the profitability of new research at risk.

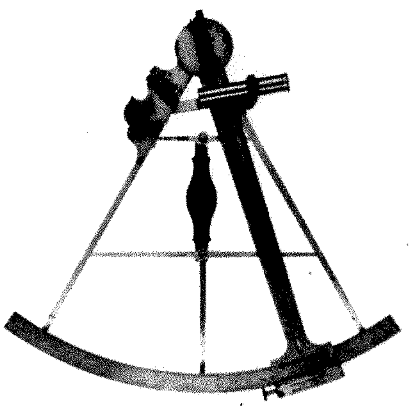
Although the European Convention on patents was signed eight years ago by ten European countries, three (Ireland, Denmark and Greece) have yet to ratify it. Italy has been particularly heavily criticized. Although it has ratified the convention, a new bill on drug patenting is being debated in the Italian parliament which appears to conflict with the convention. It is claimed that the Italian Senate has amended the bill to give protection to Italian firms manufacturing drugs which elsewhere would be protected by patent law. And in Yugoslavia manufacturers are said to be exporting low price drugs using patents provided by European countries solely for use in the domestic market. **Jasper Becker**

Cook's aid surfaces

A 15-inch radius brass sextant which is thought to have been taken on the last of Captain James Cook's historic voyages of exploration (1776–80), was sold at Sotheby's on 11 June for £11,000; it was made by Jesse Ramsden around 1775.

Ramsden was born in Halifax, Yorkshire, in 1735. His mathematical studies were interrupted by his father apprenticing him to a cloth-worker in Halifax; he then worked until about 1758 in a cloth-warehouse in London. It was here that he became apprenticed to a relatively unknown instrument maker named Barton (or Burton) in Denmark Court, Strand.

In 1763 Ramsden set up business on his own account in Haymarket and in 1775 he moved to Piccadilly, London, where he remained until his death in 1800. From the outset of his business life, his skill as an engraver and divider attracted the attention of leading London makers such as Nairne, Sisson, Adams and Dollond.



His design and construction of one machine revolutionized the dividing of the scales of sextants and other instruments. The Board of Longitude (set up by the Admiralty in 1714) awarded Ramsden £615 on the understanding that he would undertake to divide octant and sextant scales for other makers.

There is little doubt that it was directly due to the inventive genius and practical skills of Ramsden that the sextant was improved to the point that an accuracy of 10 or 20 seconds rather than 1 minute became the order of the day. In addition, the finely engraved scales made possible by the use of the dividing machine permitted sextants to be made smaller than previously without loss of accuracy.

Ramsden's improvements to the sextant eased life for the sailor, for latitude could now be determined with an unprecedented degree of accuracy. Even longitude, which had seemed to present insurmountable problems, was in the process of being solved, the solution being the provision of a suitable timekeeper, an accurate sextant and a nautical almanac. **Arthur Frank**

0028-0836/81/250529-01$01.00

CORRESPONDENCE

Role of US Academy

SIR — Your editorial of 30 April (page 723) commenting on the recent history of the National Academy of Sciences under the leadership of Philip Handler was generous indeed. Nevertheless, it calls for some supplementation, lest your readers carry away a seriously deficient understanding of this institution's present concerns. The editorial urges the Academy to undertake study of a range of problems, both old and new, when — in fact — such studies are long underway. Some have already been completed.

"The education of students in schools and colleges?" or "The problem of university scientists, with or without tenure, engaged in the hazardous gamble of the pursuit of scholarship?". Our Commission on Human Resources has provided major studies of both these areas to the Federal Government within the past year.

"The relationships that should obtain between universities and the federal government with respect to the support and conduct of research? The impact on education of the new strict interpretation of rules prohibiting the export of military material?" The appointment of blue-ribbon committees to examine all aspects of these issues and to report to the Council of the Academy was announced several months ago.

In addition, a committee is being formed to look into the propriety of the arrangements academics make with universities to share in the profits of extramural commercial enterprises. With regard to new regulations from the Office of Management and Budget on time-accounting, the officers of the Academy and the membership at large have taken several steps to present their views to appropriate government officials, both publicly and privately.

Most baffling to us is the final question, "whether . . . to protest at Sakharov's exile and, if so, how?". Although many individuals and institutions have registered strong protests against the official harrassment of Andrei Sakharov, surely the National Academy was among the first to express its grave concern over the treatment of its distinguished foreign associate — first privately and, since 1973, most publicly. Following Sakharov's internal exile, the Academy Council cancelled a symposium planned jointly with the Soviet Academy and deferred all other such symposia for a period of six months. With Sakharov's status unchanged, the Academy Council reaffirmed that position in August 1980 and just over a month ago.

Permit me one further observation. No one could disagree with your observation that the Academy should arrange to "shuffle off commissions for studies which are either pointless or untimely". Indeed, to reduce that likelihood, President Handler reorganized the National Research Council early in his first term to insure that all new activities of the Research Council had first been evaluated and authorized by appropriate bodies composed primarily of Academy members. Even so, there will always be disagreement over the relative significance of new studies. What is significant to A may appear quite trivial to B. For example, if your editorial meant to cite as pointless and untimely the study of chlorofluouromethanes and the ozone layer, and of the health effects of dietary cholesterol and low-level ionizing radiation, Dr Handler, I know, would strongly disagree.

Even more fundamental is our belief that in a pluralistic society an institution is not likely to be permitted to answer only important questions.

HOWARD J. LEWIS

Director, Office of Information,
National Academy of Sciences,
Washington, D.C., USA

Questionable results

SIR — In answer to your editorial question "Who, not Congress, should police fraud?"[1], a recent proposal by Donald Goodwin[2] deserves consideration. He suggested that a Committee for Replication composed of scientists should be established. Each year the committee would select one or two findings they deemed important and surprising enough to warrant a special type of replication: replication at the laboratory from which the original finding came, but with an independent investigator present as a researcher, witness, and reporter. The independent investigator would be a scientist chosen by the committee from within the same field of research but from a different laboratory.

Goodwin's proposal has several merits:

(1) The process is directed towards establishing the reliability of dramatic new findings — the primary concren for scientists — and not specifically towards the discovery of fraud. Incorrect findings may occur in print for a variety of reasons, probably the rarest of which is fraud; the procedure would help expose incorrect findings regardless of cause.

(2) The process is more accurate in some ways than replication at other laboratories: for example, there may be critical differences between the procedures used at different laboratories that are not clear from the published accounts. Furthermore, in many fields it is rare that a pure replication is attempted; usually there are modifications, "improvements", introduced by the other laboratories. There has been generally only a low reward for conducting replications and perhaps as a consequence, many experiments were not duplicated[3]. The ratio of reward to effort for the independent investigator would probably be much higher with Goodwin's procedure, which also would assure the replication of the important experiments the committee selected.

(3) As mentioned in your editorial[1] "the scales are already too powerfully weighted against the heterodox". This procedure would not add to the weight against surprising findings and might help to correct the balance. Confirmation of a dramatic result by on-site replication would increase its acceptability and help to communicate the finding to a wider audience. It would also usually produce faster replications than relying on other laboratories.

(4) The process avoids the stigma of the witch-hunt. It would be an honour to have one's research choosen for on-site replication, a demonstration that the scientific community recognizes the great importance of the results. It would provide an opportunity for an honest researcher to show the reliability of the results to sceptics. Nevertheless, the procedure would still act as a deterrent to fraud.

(5) Although the process would need to be funded, probably by a government agency, the actions would still be controlled by scientists. It would still be in the realm of peer review.

Such on-site replication is not, of course, a panacea, and there are several difficulties associated with it. From my own experiences as the independent investigator in the on-site replication attempt[4] that initially stimulated Goodwin's proposal, I am quite aware of some of these dificulties, but I also can vouch for the overall feasibility of the procedure.

J.D. SINCLAIR

Research Laboratories of the State Alcohol Monopoly (Alko), Helsinki, Finland

1. *Nature* **290**, 433–444 (1981).
2. Goodwin, D.W. at *XII Annual Nordic Meeting on Biological Alcohol Research*, April 21–23, 1981 Stockholm, Sweden.
3. Broad, W.J. *Science* **212**, 421 (1981).
4. Sinclair, J.D. at *XII Annual Nordic Meeting on Biological Alcohol Research*, April 21–23, 1981.

Film loop evolves

SIR — In his review of the Natural History Museum's new exhibition, *Origin of Species*, Professor Cox (*Nature* 4 June, p.373) refers to a "film loop" dealing with the status of the theory of evolution by natural selection. He regards this audio-visual programme as a failure and, apparently, misconceived.

In producing this audio-visual we were trying to avoid dogmatism in our presentation of the theory of evolution by natural selection. The intention was to reproduce the arguments set out in Chapter 12 of Colin Patterson's book, *Evolution* (British Museum (Natural History), 1978). Before the opening of the exhibition the Museum was aware that as a consequence of compression of the subject matter and transference to a new medium the audio-visual might give an impression other than that intended. Subsequent experience of the audio-visual as part of the exhibition has shown this misgiving to have been justified. A new version is being prepared.

R.S. MILES

Department of Public Services,
British Museum (Natural History),
London SW7, UK

Had enough too?

SIR — It is high time that you terminated the correspondence on the "philosophical" status of evolution. If some learned and pompous ass asserts that Darwin's theory is not "really" scientific, or that it is "socially embedded" or that holders of conflicting theories cannot "really" talk to one another about the same phenomena, then the necessary and wholly sufficient reply is one good old word that you probably wouldn't print.

M. HAMMERTON

Department of Experimental Psychology,
University of Oxford, UK

0028-0836/81/250530-01$01.00

NEWS AND VIEWS

Gene transfer in plants

from Roy Davies

PLANT BREEDERS often need to introduce a single character from one genotype into another; a classic example is the introduction of a different gene or allele for disease resistance into an existing high-yielding cereal variety to combat an emergent virulent strain of a pathogen. The usual way of doing this is by recurrent backcrossing, but the generation times of higher plants makes this a slow and laborious procedure. In this issue of *Nature* (p.586), Jinks and his colleagues at the University of Birmingham describe an alternative and more rapid method of achieving this transfer of one or a few attribute(s) from one genotype to another.

Two diploid inbred lines (V27 and V12) of *Nicotiana rustica* were used in the experiments, with V27 being used as the maternal parent in the crosses and pollen of the V12 parent being exposed to doses of 0, 10, 15 or 20 krad of γ radiation. The two genotypes differ in several major genes; V27 plants are yellow, have a mop inflorescence and yellow ovaries, each of these characters being controlled by more than one recessive gene. V12 plants are green, have a non-mop inflorescence and black ovaries. In addition, the two lines are distinguishable in terms of several quantitative characters. For all the attributes listed, V12 has the dominant and V27 the recessive genes. As expected, the first generation plants derived from the crosses involving irradiated pollen (M_1 generation) were more variable than the F_1. Plants of the M_2 generation, while still showing some variation, resembled the maternal parent, V27, more than did the F_2, both for those characters determined by major genes and for the continuously variable characters. This resemblance increased with dose administered to the pollen. M_2's resemblance to V27 is contrary to expectation since the pollen parent V12 contained the dominant alleles for the characters studied.

Up to this point the data could be interpreted in terms of a radiation-induced loss or mutation of the dominant alleles of the V12 parent, thus making the offspring more like the parent having the recessive genes, V27. However, it is the other features of the data that make the experimental results of particular interest. Of 97 M_2 plants derived from the 20 krad treatment, 16 had the yellow plant colour of V27 and 10 were also mophead — a combination not found in the 100 F_2 plants examined. (There are several genes involved in the regulation of each of these characters.) Of these 10 plants resembling the maternal parent, that is, which were yellow mop, 6 had black ovaries — a paternally derived attribute. In terms of the three quantitative characters scored, all 16 yellow plants were very similar to the maternal parent. From previous experiments (Virk *et al. Heredity* **39**; 287, 1977) the authors are confident that true-breeding yellow, mop, black-ovaried genotypes can be generated in later generations. Thus from a small M_2 generation, plants could be isolated which were similar to the maternal parent, V27, for the two characters controlled by major genes, and for the three quantitative characters scored, but incorporating a single attribute from the paternal parent. Again an M_2 plant similar in all respects to V27 apart from one V12-like quantitative character was identified.

Comparable observations have been reported by Pandey in other species of *Nicotiana* (*Nature* **256**; 310, 1975; *Heredity* **45**; 15, 1980). He succeeded in making analagous transfers of particular characters both within and between species after exposure of pollen to ionizing radiations, albeit using much higher doses even than those used by Jinks and his colleagues.

These two sets of results are of considerable interest both because of their practical implication for plant breeding and for the mechanism involved. There is now a real incentive for plant breeders to examine the extent to which this technique can be used for species other than those belonging to the genus *Nicotiana*, not only for the more rapid transfer of characters in intra-specific crosses, but also for the transfer of characters in those inter-specific or inter-generic crosses where, due to the incompatibility of the two genomes involved, the embryo aborts early in development. The introduction of only one or a few characters from another species or genus, rather than a whole alien genome, may well be tolerated, and allow the normal development of an embryo. There is yet another avenue to be explored; the production of somatic hybrids between species and genera of plants by the fusion of protoplasts is a rapidly developing technology. The hybrids themselves are unlikely to be directly useful as new crop plants, but by simulating the experimental approach used by Jinks and Pandey, and irradiating one of the parental protoplasts before fusion with an unirradiated 'recipient', specific desirable attributes could be transferred. A recent paper by Dudits *et al.* (*Molec. gen. Genet.* **179**; 283, 1980) illustrates a first step in this direction. X ray-irradiated protoplasts of parsley ($2n = 22$) were fused with unirradiated carrot protoplasts ($2n = 18$) and among the regenerated plants selected were some which had 19 chromosomes; these had some genetic markers from the irradiated donor.

What are the mechanisms involved? Pandey has shown that his 'transformed' plants are diploid and show no gross chromosomal abnormality; furthermore, where the two parents differed in chromosome number the 'transformed' offspring had the maternal parent's chromosome number. In the present *N. rustica* experiments, cytological studies have not been undertaken but the fertility and morphology of the plants indicates that they are neither haploid nor grossly aneuploid. It appears that irradiation of the pollen fragments the chromosomes, but pollen tube growth can still occur even after the high radiation doses. The

---

Roy Davies is Professor of Applied Genetics at the University of East Anglia, Norwich.

0028-0836/81/250531-02$01.00

fragments are transferred to the egg and parthenogenetic development of the egg occurs. The chromosome number of the egg or early embryo then doubles and finally some of the fragments are incorporated into, or associated with, the maternal genome. A crucial point to be investigated is whether the chromosome doubling and parthenogenetic development that occur so readily in these *Nicotiana* species, can occur as readily in other genera.

The size and nature of the chromosome fragments from the pollen parent, and the mechanism of their integration or association is of great interest. Genetical studies of linked markers and of segregation patterns in later generations will help to determine the size and location of the associated or integrated fragments, and physical characterization of particular DNA fragments will be feasible. A general assumption of radiation cytogenetics is that broken ends of chromosomes rejoin only with other broken ends; since the maternal parent is unirradiated such ends will not be available for rejoining with the fragments produced in the pollen. How else might integration occur, if this is indeed what happens? Fragments could be incorporated during replication, or during the exchange events that ultimately become recognized as sister chromatid exchanges; if this is so, then the average size of the fragments would be expected to be small. Transposable elements of the kind first recognized in McClintock's work on maize, and physically characterized, for example, in *Copia* in *Drosophila,* must have a chromosomal organization which allows integration into, and excision from, eukaryotic chromosomes; but it seems unlikely that random fragments induced by radiation would have a similar organization. One of many puzzles posed by the *Nicotiana* results is the frequency with which transfer of attributes occurs; 6 out of 97 M_2 progeny generated after a 20 krad dose of γ radiation to the pollen closely resembled the maternal parent yet had the black ovary character from the paternal parent. It remains to be seen whether this is stable in later generations and whether black ovary is typical in its rate of transfer; if it is, then this represents a high rate of integration or association. Clearly there are a large number of issues to be resolved but these investigations by Pandey and by Jinks and his colleagues should now provide an incentive for a great number of workers to re-examine the basis and practical application of the phenomenon of gene transfer. □

Cell sorting, cell differentiation and cyclic AMP

from Adrian Tsang

IN the slime mold *Dictyostelium discoideum*, cyclic AMP is now believed to be directly involved in the processes of cell sorting and differentiation, as well as acting as a chemotactic agent in mediating aggregation. This is one of the conclusions that emerged from a recent meeting* on the development of cellular slime molds.

Amoebae of *D. discoideum* multiply as isolated cells. The developmental programme is set in motion at the onset of starvation, and after about six hours the amoebae aggregate in response to cyclic AMP signals. The multicellular structures, called slugs, which are formed after aggregation are differentiated along the anterior–posterior axis. The anterior, prestalk region gives rise to the stalk cells of the terminally differentiated fruiting body, while the cells of the posterior region are the precursors of the mature spores. This formation of a patterned arrangement of new cell types has obvious analogues in the development of higher eukaryotes.

It has been known for some years that prestalk and prespore cells can sort out, but it is not entirely clear whether sorting out precedes differentiation or is a consequence of it (Forman and Garrod *J. Embryol. exp. Morph.* **40**; 229, 1977). The answer to this question is important because it will indicate whether or not the position a cell occupies in the aggregate has an effect on its differentiation. Making use of the observation that amoebae grown in a non-glucose medium, when mixed with amoebae grown in a glucose-containing medium, differentiate preferentially into prestalk cells (Leach *et al. J. Embryol. exp. Morph.* **29**; 647, 1973), Takeuchi (Kyoto University) examined the relationship between sorting out and differentiation. He mixed ^{3}H-labelled, non-glucose-grown amoebae with glucose-grown amoebae and allowed them to form agglomerates in submerged conditions. At various times the agglomerates were sectioned and stained with anti-spore serum to locate the prespore cells, and the same sections were autoradiographed to locate non-glucose-grown cells. Takeuchi found that the prespore cells began to differentiate long before the labelled and unlabelled cells were sorted out, suggesting that cell sorting is not a prerequisite for differentiation. This thesis is supported by a separate experiment performed by Okamoto (Kyoto University) in which dissociated aggregative cells were shaken at high speed. Under these conditions, small clumps of cells were formed and the prespore cells in the clumps were found to be randomly distributed. As no cell sorting can occur in these small clumps of cells, the above data indicate that *D. discoideum* cells can differentiate in the absence of cell sorting. In other words, whether a cell differentiates into a prestalk or a prespore cell may not depend on the position it occupies in the multicellular aggregate.

Cellular sorting out has often been attributed to differential adhesion. That this may not be the whole story in *D. discoideum* was demonstrated by Sternfeld (Albert Einstein College) and Takeuchi. Sternfeld found that when dissociated slug cells were allowed to form clumps in suspension, a new pattern was formed with the prestalk cells in the centre of the clump. However, if the culture medium contained cyclic AMP, the prestalk cells were found on the periphery. Moreover, Takeuchi showed that prestalk cells possess higher motive force than do prespore cells (Inouye and Takeuchi *J. Cell Sci.* **41**; 53, 1980). Although these experiments do not rule out the possible participation of differential adhesion in sorting out, they do strongly support the earlier suggestion of Matsukama and Durston (*J. Embryol. exp. Morph.* **50**; 243, 1979) that cell sorting is guided by differential chemotaxis to cyclic AMP.

Evidence accumulated over the past decade indicates that cyclic AMP participates in the differentiation of stalk and spore cells. Both Sussman (University of Pittsburgh) and Rutherford (Virginia Polytechnic Institute) showed that the potential of establishing a cyclic AMP gradient exists in the multicellular aggregate. Blumberg (MIT) and Okamoto provided additional evidence that the accumulation and the maintenance of post-aggregative gene products require cyclic AMP (Town and Gross *Devl Biol.* **63**; 412, 1978), and there is direct evidence that cyclic AMP affects the expression of developmental gene products by altering the rates of transcription of specific genes (Williams *et al. Proc. natn. Acad. Sci. U.S.A.* **77**; 7171, 1980). What remains obscure is the role of cyclic AMP in cell differentiation. Is it a general effector of gene expression, a morphogen which channels the cells into specific differentiative pathways, a maturation factor for stalk and spore cells or a combination of these? □

Adrian Tsang is at the Imperial Cancer Research Fund Mill Hill Laboratories, London, UK.

---

*The US-Japanese conference on the development of cellular slime molds held in San Diego, California in February was organized by I. Takeuchi (Kyoto University) and D. Francis (University of Delaware).

0028-0836/81/250532-01$01.00

Scratches on the Palaeolithic record

from C.S. Gamble

THE study of our earliest ancestors often reveals a preoccupation with thoughts of food. In particular, the importance of meat in the diet of early hominids and the means by which it was obtained and shared has led to conflicting accounts about the cooperative[1] and aggressive[2] origins of the human species.

The difficulty of deciding between these alternative pictures of the past suggests that occasionally our understanding of early man owes more to our contemporary imaginations than to our ability to interpret the archaeological record. Current research is now tackling on an interdisciplinary front the complex issues of how this record is formed and how we can infer patterns of past behaviour from such contemporary observations.

An example of this approach is to be found in the continuing investigation of the early man sites at Olduvai Gorge, Tanzania. The studies reported in this issue of *Nature* by Bunn (p.574) and by Potts and Shipman (p.577) show how the well known data from these sites[3] can still provide surprises when squeezed a little harder. In her landmark report, Leakey described many of the excavated sites at Olduvai as 'living floors'. This term implies that the bones and chipped stones found there were in some way causally related and that these observations could be used to reconstruct patterns of hominid behaviour that involved tool use, and its relationship to dietary behaviour.

But what if these and other East African sites were merely hydraulic jumbles[4] where the association of stone tools with animal bones was the result of natural agencies, such as river transport, rather than hominid activity? Moreover, even if these natural agencies could be discounted, how could we distinguish between the bones that resulted from carnivore as opposed to human food refuse? The new observations by Potts, Shipman and Bunn provide good evidence that at least some of the bones and chipped stone tools were found together as a result of hominid rather than hydraulic activity. They have identified cut marks on the surface of herbivore long bones, which can be shown by use of a scanning electron microscope to correspond with those produced under experimental conditions by stone tools. The experimental work classifies various scratches as cut, slicing, scraping and chop marks, which can be distinguished from carnivore chewing, gnawing and bone crushing and from the distinctive gnawing patterns of rodents such as the porcupine. Scratches left on bones by carnivore claw marks have not yet been examined.

Most of these cut marks have been found superimposed over carnivore scratches, but in at least one case this relationship is reversed, indicating that at least once a hominid got to the bone before the lions, hyenas and other carnivores of 1.8 million years ago. At the famous Zinj site at Olduvai, Bunn found that, from among a total of 3,500 animal bone fragments, about 300 bore cut marks while a further 400 were chewed by carnivores. Additional evidence comes from the recent excavations at Site 50, Koobi Fora in Kenya[5]. Here it was possible to fit together bone fragments, several of which also bore cut marks. The fragmentation of these long bones appeared in some cases to be related to the cracking of bones for marrow. Cut marks have also been found on bones which come from excavated sites where stone tools are not present.

These findings undoubtedly add a further dimension to our understanding of the archaeological record of early man. First, the behavioural association between stone tools and animal bones seems substantiated for at least part of these excavated assemblages. Second, the closer examination of carnivore gnawing on bones points to the composite nature of these sites and the fact that archaeologists must continue to investigate ways of distinguishing between the archaeological and palaeontological records[6]. These observations of cut marks are also important because of the difficulties of analysing traces of wear on stone tools. The recognition of distinctive polishes on the edges of flint artefacts which result from cutting and working meat, hides, wood, bone and plant material[7] has opened up many research possibilities, but this is not yet possible for the lava, quartz and quarzite tools that make up the majority of tools from Olduvai and Koobi Fora. The proposal that these tools were used for stripping meat from carcasses and as hammers for cracking bones for marrow rests on the observation of cut marks and bone fracture patterns. Supporting evidence for the dietary practices of early man may also come from scanning electron microscope studies of fossil hominid teeth. Studies by Walker[8] of baboons and hyraxes has shown that different diets lead to different attritional patterns on tooth surfaces and these investigations are now being applied to specimens of fossil man[9].

What do these observations contribute to our understanding of the behaviour of early man? Although they show that meat and marrow were processed and consumed, they do not directly answer the question of how, through either hunting or scavenging, this food was obtained.

This is a problem of inference. How can archaeologists reconstruct the dynamics of past hominid behaviour when they can only work on the inert stone tools and bones that they dig up? This problem is further complicated in the study of early man since there are no contemporary examples providing suitable behavioural models to help interpret the excavated archaeological patterns. We can argue that the association of stone tools and animal bones demonstrates that meat was eaten and suggest that it was acquired by hunting or scavenging; but although it is a reasonable explanation for such archaeological observations, we are still mostly guessing at what went on in the past.

L.R. Binford[10] considers that one of the commonest residues found on archaeological sites, the structure of animal skeletons, provides a constant against which some aspects of hominid and carnivore behaviour can be measured. Observations of skeletal disarticulation and bone destruction by carnivores and contemporary hunter-gatherers[11] reveal very different patterns that can be related to uniformitarian principles of the mechanics of feeding. Using the work of A.P. Hill on the taphonomy of East African large ungulates and his own observations, Binford shows that it is possible to distinguish between residual assemblages of animal bones that have been primarily utilized by man and by carnivores. Moreover, he shows how these differences can be accounted for in terms of a predictable set of variable behaviour. In other words, it is possible to take a collection of inert observations from the archaeological record and infer the dynamic pattern of past behaviour that resulted in those residues.

Binford has applied this methodology to the 'living floors' from Olduvai Gorge. He concludes that the anatomical parts of the large animals that were found in these sites fit the model of carnivore feeding and that we are looking at ravaged carcasses. Not all of the patterning from these animal bone collections could, however, be explained

---

C.S. Gamble is in the Department of Archaeology, University of Southampton.

---

1. Leakey, R. & Lewin, R. *People of the Lake* (Avon, New York, 1978).
2. Ardrey, R. *The Hunting Hypothesis* (Atheneum, New York, 1976).
3. Leakey, M. *Olduvai Gorge* Vol.3 (Cambridge University Press, 1971).
4. Binford, L.R. *J. anthrop. Res.* **33** (1977).
5. Bunn *et al. World Arch.* **12** (1980).
6. Behrensmeyer, A.K. & Hill, A.P. (eds) *Fossils in the Making* (University of Chicago Press, 1980).
7. Keeley, L. *Experimental Determination of Stone Tool Uses* (University of Chicago Press, 1980).
8. Walker, A. *Fossils in the Making* (eds Behrensmeyer, A.K. & Hill, A.P.) (University of Chicago Press, 1980).
9. Walker, A. *Phil. Trans. R.Soc.* **B292,** 57 (1981).
10. Binford, L.R. *Bones: Ancient Men and Modern Myths* (Academic, New York, 1981).
11. Burford, L.R. *Nunamiut Ethnoarchaeology* (Academic, New York, 1978).

0028-0836/81/250533-02$01.00

by the carnivore model — some marrow bones which should have been present were missing and the differential pattern of bone destruction favoured additional agencies involved in splitting bones for marrow. This is the information about hominid behaviour. These hominids were scavenging the kills of large predators for bones of low food utility, scraping off the last scraps of meat and cracking them for marrow which the carnivores had ignored. Transport of the scavenged bones away from the kill site and to more protected locations also took place. Stone tools to crack open these bones and to slice off the last pieces of meat gave man a particular advantage in competing for this niche that utilized the waste of the predators. The evidence for cut marks and the investigation of bone breakage patterns would seem to support this reconstruction. The strong inferential base of this behavioural reconstruction now allows us to measure the activities of early man against the evidence from the rich palaeontological record of which the archaeological evidence forms a small but significant part.

The discovery of free quarks?

from Louis Lyons

SINCE earliest times there has been speculation on the nature of the ultimate constituents of matter which are responsible for the enormous variety of objects that exist in the Universe. By the beginning of this century, it was believed that there were just two 'elementary particles' — the proton and the electron — out of which everything was constituted. This satisfactory state of affairs was gradually eroded with the discovery of more and more particles, so that today the list of so-called elementary particles contains several hundred entries. There thus exist more 'elementary particles' than chemical elements.

In 1964, Gell-Mann (*Phys. Lett.* **8**; 214, 1964) and Zweig (CERN preprint TH412) independently suggested that the multitude of elementary particles were made out of simpler objects called quarks. At the time only three quarks were needed to explain the existence and many of the properties of all known particles, as well as the absence of certain conceivable types of particle which have not been found. Subsequently the quark model was also used to explain many of the features of scattering experiments involving the composite 'elementary particles' (Feynman *Photon–Hadron Interactions* Benjamin, New York, 1972).

The simplicity and success of the quark model immediately lead to a succession of experiments designed to detect free quarks. Despite the fact that all elementary particles have electric charges which are integers (when measured in units of e, the electron's charge), most versions of the quark model require quarks to possess fractional charges of ±e/3 or ±2e/3. Free quarks should also possess masses different from the conventional elementary particles. Experimental observations of unusual charges and/or masses could be evidence for the existence of free quarks.

Despite the variety of tested materials (which have included the Sun, moon soil, meteorites, ocean sludge, mountain lava, oysters, cosmic rays and particles produced in high-energy interactions at accelerators), most experiments have provided no convincing evidence for the existence of free quarks. Thus Morpurgo and his collaborators in Genoa have used a modification of the original Millikan oil drop experiment to measure the charge on steel balls. They find no fractional charges and set a limit of less than 1 quark in 10^{21} nucleons.

Other experiments have set even lower upper bounds on quark concentrations, but such experiments have involved quark enrichment and transportation techniques that rely on assumed chemical and physical properties of quarks. A positive signal from such experiments would have been significant, but the negative results may simply mean that free quarks do not behave in the assumed way. Quark search experiments have recently been reviewed by Jones (*Rev. Mod. Phys.* **49**; 717, 1977) and by Lyons (*Prog. Particle nucl. Phys.* **7**; 157, 1980).

The almost universal absence of an observed signal of free quarks leads to the concept of 'confinement': quarks are able to exist in pairs and in triplets inside elementary particles, but it is impossible (according to this hypothesis) to pull them apart and observe them singly. Perhaps a quark is like one end of a piece of string: no matter how hard one tries, it is impossible to isolate a single end. Vigorous attempts are being made to prove that confinement is a necessary consequence of Quantum Chromodynamics (QCD), a theory of the interactions of quarks with each other. Despite the lack of success to date, confinement is still held as a popular belief (Johnson *Scient. Am.* **241**; 100, 1979).

Recently, however, a Stanford experiment, using a different variant of the Millikan experiment, has reported a positive signal of fractional charges in niobium spheres of radius 1/10 millimetre (LaRue *et al. Phys. Rev. Lett.* **15**; 967, 1981). Of the 39 measurements on the 13 spheres tested, 14 values of fractional charge are observed on 5 different spheres. Furthermore, all 39 measurements give values of the charge consistent with zero or ±e/3. (In these experiments, charges of ±2e/3 are indistinguishable from ±e/3). Does this mean that free single quarks have been observed? If so their density in niobium is at least 1 quark per 10^{20} nucleons. Would this be in conflict with earlier null measurements?

There still exists the possibility that, since fractional charges have been seen in only one experimental apparatus, they are in some, not understood way an instrumental effect. For example, it is necessary to apply corrections of magnitude larger than e to the observed measurements in order to deduce the true values of the charge, and also some measurements are discarded due to apparatus instabilities during the course of a run. In order to eliminate the possibility that such effects could enable the experimenters (subconsciously) to adjust their final measured charges to be ±e/3, future measurements by the Stanford group will have a random extra fractional charge inserted into their data, and whose magnitude is unknown to them. Only after they have decided whether to accept a measurement and have completed calculating their corrections will the random charge be subtracted. If they then still find values consistent with only zero and ±e/3, the evidence that the fractional charges are genuine will be much more convincing.

The Stanford experiment's positive result is not necessarily inconsistent with the many previous null results because there are no other direct measurements on niobium to date. As far as searches in other materials are concerned, there is no particular reason to suppose that the quark/nucleon density is independent of material. However, Morpurgo points out that the Stanford spheres change their charge quite readily; this is interpreted as quarks dropping off the surface from time to time, so they should be found in other materials as well. Clearly it is desirable to repeat the search for fractional charges in niobium by as direct a technique as possible in another apparatus.

It is also possible that the Stanford fractional charges are nothing to do with the quarks that are responsible for building elementary particles, but correspond to something else altogether; indeed the Stanford paper makes no mention of the word 'quark'. If, on the other hand, free quarks have been seen, it would significantly modify current theoretical ideas about the way quarks interact with each other. Meanwhile the experimental situation needs clarification. The outcome is awaited with great interest.

Louis Lyons is in the Nuclear Physics Laboratory, Oxford.

0028-0836/81/250534-01$01.00

Raman: an eye for combustion studies

from D.A. Greenhalgh

Accurate measurements of the events taking place in the very hostile environments inside combustors such as furnaces, gas turbines and internal combustion engines has not been possible until very recently. Now, a new generation of optical measurement techniques based on the laser are offering real solutions to the problems of making remote, non-intrusive measurements.

By far the most successful technique for temperature and specific species measurements is laser Raman spectroscopy; in particular, Coherent Anti-Stokes Raman Scattering (CARS) which has been used for measurement inside internal combustion engines and jet combustors. Together with laser anemometry, which provides information on gas velocities and turbulence, these techniques could well revolutionize our knowledge of combustion processes over the next decade.

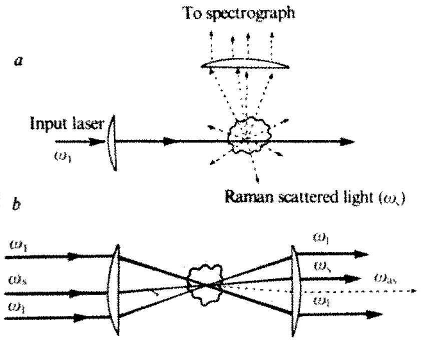


Fig. 1 *a*, Spontaneous Raman. $\omega_1-\omega_s=\omega_R$; ω_R = frequency of Raman active transition. *b*, CARS. $\omega_1-\omega_s=\omega_R$ and $\omega_1-\omega_s+\omega_1=\omega_{as}$; the 'laser-like' CARS signal is at frequency ω_{as}.

In order to appreciate the role of Raman spectroscopy in combustion, it is necessary to outline the basics of a practical combustor. The fuel is frequently a liquid hydrocarbon which is sprayed or directly injected into an air flow, and burns after a suitable mixture of vaporized fuel and air has formed. The process is often described as 'diffusion burning' and the chemical kinetics are fast compared with the mixing processes. This is, however, a rather simplified picture — even before actual combustion, the fuel molecules will be pyrolysing to form lighter molecules and soot, or its precursors.

Coal combustion is considerably more complex but nevertheless it can be modelled in a predominantly aerodynamic manner, concentrating on parameters such as air flow and temperature. To predict these parameters accurately a moderate amount of chemistry has to be included in the models. Whereas flame chemistry can be studied through measurements of well controlled laboratory flames, the physics of diffusion flames often requires information from actual devices or close replicas. Although Raman spectroscopy is important in measurement of both physical and chemical properties, the role of CARS for measurements in practical combustion flows will be emphasized here.

Raman scattering involves inelastic scattering of light by molecules. In conventional Raman spectroscopy an intense beam of monochromatic light is focused at the desired measurement point within the sample gas and the inelastic scattered light, which is Stokes-shifted from the incident light, is subsequently analysed with a spectrograph or similar instrument. Such a system only collects light from the focus of a laser beam and thereby ensures very high spatial resolution. From the spectra of simple molecular species, such as diatomics, water and carbon dioxide, it is possible to deduce both temperature and concentration of the individual component molecules. Unfortunately, conventional Raman spectroscopy is limited to studies of premixed flames or to diffusion flames, such as hydrogen and air, where soot formation is impossible. This restriction is not set by the luminosity of sooty flames but because the fluorescence induced by the incident laser light swamps the weaker Raman-scattered light.

New Raman techniques, of which CARS is the most successful, developed from research in non-linear optics, particularly that of Taran and his co-workers[2]. In conventional Raman the scattered light obeys the relationship $\omega_1-\omega_{sj}=\omega_{Rj}$ where ω_1 is the frequency of the incident light and ω_{sj} is the *j*th component of the Stokes-shifted Raman signal corresponding to each ω_{Rj}, the frequency of the *j*th individual Raman active transition. To obtain a CARS spectrum, three laser frequencies are introduced into the medium such that the signal at frequency ω_{as} (as = anti-Stokes) obeys the relationship $\omega_{as}=\omega_1+\omega_2-\omega_s$. Usually ω_1 and ω_2 are the same laser and we have the degenerate case where $\omega_{as}=2\omega_1-\omega_s$. The latter process is represented diagramatically in Fig. 1 where a comparison is drawn with conventional or spontaneous Raman. If $\omega_1-\omega_s=\omega_{Rj}$, then the process is resonant and a strong 'laser-like' emission is produced at ω_{as}. In common with conventional Raman, very high spatial resolution (∿ 2 mm x 0.2 mm diameter) is possible. In conventional Raman the incident laser excites the molecules of the medium which then spontaneously decay, releasing the Stokes-shifted signal at ω_s. Due to the spontaneous nature of the process, the signal is scattered into 4 π steradians. By comparison, CARS is a coherent process, where the laser fields at ω_1 and ω_s mix within the medium to induce a coherent signal at ω_{as}; thus the signal has the optical properties of the input lasers and is a 'laser-like' beam. This simple facet of CARS is its main strength; a laser beam can be collected efficiently by optics with a very small aperture whereas laser-induced fluorescence or a conventional Raman signal cannot. This advantage has been realized for several years and was demonstrated in sooty flames by Beattie and co-workers[4].

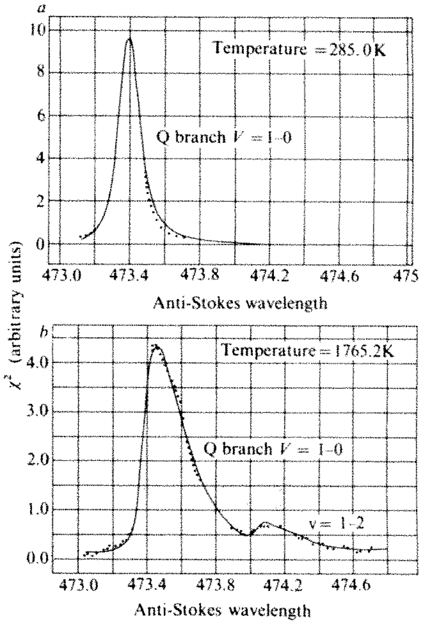


Fig. 2 χ^2 plot. *a*, Room temperature N_2 spectrum. *b*, 'Single-shot' (10^{-8}s) N_2 spectrum from Gem combustor can.

There is another fundamental difference between the two techniques: in conventional Raman a complete spectrum is generated by one monochromatic incident wavelength ω_1 whereas only one frequency component, ω_{as}, of the CARS spectrum can be generated if ω_1 and ω_s are monochromatic. In order to generate a Raman spectrum by CARS, one laser, normally ω_s, is frequency tuned so that a spectrum of the Raman active transitions is produced. This approach has been used to demonstrate CARS in both a gas- and a liquid-fuelled internal combustion engine by Stenhouse *et al.*[2], and in a combustion tunnel burning kerosene fuel, by Taran[6]. Although both these experiments demonstrated the feasibility of temperature measurements from the Raman

---

D.A. Greenhalgh is in the Engineering Sciences Division, AERE Harwell, UK

0028-0836/81/250535-02$01.00

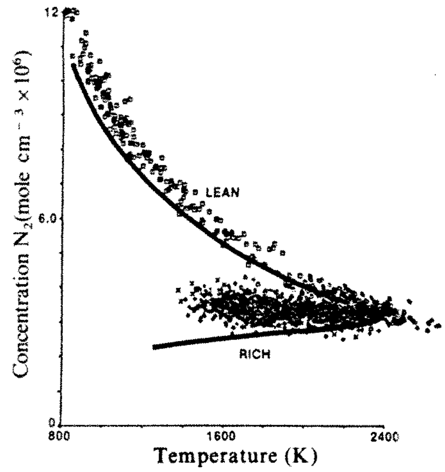


Fig. 3 Comparison of simultaneously acquired N_2 concentration and temperature data with a theoretical curve for a hydrogen air flame (taken from ref.9).

spectrum of nitrogen, Taran's work also demonstrated the possibility of CO_2 detection down to an estimated 1,000 p.p.m. and the measurement of CF_4 at around one per cent. The latter was injected into the combustion flow for short periods of time to infer residence times.

This approach of obtaining a spectrum by scanning the frequency of ω_s is satisfactory for demonstration experiments and laboratory gas samples, but it can lead to errors in determining mean values of temperature and concentration in fluctuating or turbulent conditions. Because the Raman spectrum is a non-linear function of temperature (it is a function of the Boltzman distribution of population of energy states), averaging varying spectra does not lead to a true mean temperature. In order to freeze these fluctuations in time, short-pulsed lasers are used (pulse length $\sim 10^{-8}$ s) and the complete spectrum is obtained in one shot. In conventional Raman this is simple since all of the spectrum is generated simultaneously, but for CARS to achieve the same result it is necessary to use a 'broad-band' laser at ω_s so that a complete section of a Raman spectrum is generated at the CARS signal frequency ω_{as}.

This approach has been very successfully developed by Eckbreth and co-workers[7] who have demonstrated single-shot broad-band CARS measurements in a variety of flames. In particular, they have measured temperature in a Pratt and Whitney JT-12 combustion can burning Jet 'A' fuel. Similar studies have recently been completed in this country on a Rolls-Royce 'Gem' combustion can fuelled with gas[8]. A typical 'single-shot' flame and a room temperature nitrogen spectra are shown in Fig.2, together with least-square fits of theoretical spectra. These spectra clearly show a large variation with temperature. In this study a sufficiently large number of 'single-shot' spectra from each geometric point was obtained to allow an estimation of both the mean and the r.m.s. of the temperature. Statistical information of this type is used to validate the aerodynamic type of mathematical model discussed earlier.

The above examples clearly demonstrate the potential for CARS spectroscopy for physical and chemical measurements in practical combustion devices, and suggest that CARS might replace conventional Raman as the preferred technique for combustion analysis. Complete replacement is unlikely since CARS is both financially and technically more demanding than conventional Raman, and the latter remains well suited to combustion measurements where signal strength and fluorescence rejection are not paramount. Indeed, fundamental combustion studies have clearly illustrated the wide-ranging ability of conventional Raman. Of particular interest is the work of Lapp and co-workers[9] where data have been collected simultaneously for temperature and nitrogen, hydrogen and water concentrations for each laser shot in a hydrogen air diffusion flame. Correlated data of this type can be very usefully compared with theoretical models. Such a comparison for a temperature–nitrogen concentration correlation is shown in Fig.3. Models of this type assume fast chemistry with the consequence that the combustion process is dominated by fuel air mixing. Although results show good

---

100 YEARS AGO

THOUGHT-READING

THE public mind has of late been somewhat agitated by the doings of a Mr Bishop, who has come before the world of London society in a capacity no less startling than that of a professed reader of thought. He has not only taken by storm the general public and daily press, but also succeeded in convening an assembly of scientific men to witness his performance, while still more recently he has had the honour of exhibiting his powers before the Heir Apparent to the Crown. It seemed to Prof. Croom Robertson worth while to make a careful trial of Mr Bishop's powers, and he therefore invited Mr Francis Galton, Prof. E.R. Lankester, and myself to join him in an investigation.

First, Mr Bishop was taken out of the room by me to the hall down stairs, where I blindfolded him with a handkerchief; and, in order to do so securely, I thrust pieces of cotton-wool beneath the handkerchief below the eyes. While I was doing this, Mr Sidgwick was hiding a small object beneath one of the several rugs in the drawing-room; it having been previously arranged that he was to choose any object he liked for this purpose, and to conceal it in any part of the drawing-room which his fancy might select. I then led Mr Bishop up stairs, and handed him over to Mr Sidgwick. Mr Bishop then took the left hand of Mr Sidgwick, placed it on his (Mr Bishop's) forehead, and requested him to think continuously of the place where the object was concealed. After standing motionless for about ten seconds Mr Bishop suddenly faced round, walked briskly with Mr Sidgwick in a direct line to the rug, stooped down, raised the corner of the rug, and picked up the object.

It was soon found that Mr Bishop succeeded much better with some of us than with others; so at a second meeting, in order to make a numerical comparison, he was requested to try two experiments with each of the four persons who were present. With Mr Galton, Prof. Robertson, and Prof. Lankester he failed utterly, while with myself he succeeded once perfectly and the second time approximately. For on the first occasion I concealed a pocket-matchbox upon the top of a book behind the leather lap of a book-shelf. After feeling along the rows of books for some time he drew out the one on which the matchbox was lying. In the second experiment I placed a visiting-card on the key-board of a grand piano and closed the cover. After going about the room in various directions for a considerable time he eventually localised the piano, and brought his finger to rest upon its upper surface about six inches from the place where the card was lying. It has also to be mentioned that in one of the experiments which he tried with Prof. Robertson the evening before, he was, after a good deal of feeling about, successful in localising a particular spot on an ordinary chair which Prof. Robertson had selected as the spot to be found. From this it will be seen that it made no difference whether a particular article or a particular spot was thought of; for if the subject thought of was a certain square inch of surface upon any table, chair, or other object in the room, Mr Bishop, in his successful experiments, would place his finger upon that spot. Neither did it make any difference whether the article or place thought of was at a high or a low elevation. Thus, for instance, in one of the experiments I placed a small pencil-case high up in the chandelier of one of the drawing-rooms. While feeling over the surface of a table in the other drawing-room, and not far from the corresponding chandelier, Mr Bishop suddely pointed at arm's length vertically to the ceiling. He remained motionless in this position for a few seconds, and then set off at a brisk pace in a straight line to the other drawing-room, until he came beneath the other chandelier.

From this brief summary it will be seen that on the whole his power of localising objects or places thought of by a person whose hand he clasps is unquestionably very striking. Of course the hypothesis which immediately suggests itself to explain the *modus operandi* is that Mr Bishop is guided by the indications unconsciously given through the muscles of his subject. He describes his own feelings as those of a dreamy abstraction or "reverie," and his finding a concealed object, etc., as due to an "impression borne in" upon him. But however this may be all our experiments have gone to show that Mr Bishop owes his success entirely to a process of interpreting, whether consciously or unconsciously, the indications involuntarily and unwittingly supplied to him by the muscles of his subjects.

from *Nature* **24**, 23 June, 171, 1881.

agreement for the lean zones of combustion, there are some discrepancies in the theory for the rich combustion zones. However, this type of approach will be of considerable value to the development of combustion models.

All the previous examples are concerned with turbulent diffusion flames where fuel air mixing is the dominant feature of the combustion process. Many of the fundamental chemical processes of combustion have been deduced from studies of premixed lamina flames, where the combustion process is stable in space and time. Conventional Raman spectroscopy can be readily applied as long as the stoichiometry is arranged for an excess of air so that soot formation is largely suppressed; for example, Bechtel and Blint[10] have examined the differences in flame composition in the main stream of a combusting flow and close to a wall. This type of study is of great interest since most practical combustors have cooled surfaces which undoubtedly affect both the physics and chemistry of the flow.

With regard to the wider applications of Raman techniques for gas-phase studies, consider the comparison between a gaseous chemical reactor and a combusting flow. In the former case the temperature is lower (∿ 800K compared with 1,700K) and the reactants are not necessarily hydrocarbon and oxygen. However, the flow is often turbulent and particle laden as in a (practical) combustor. Application of CARS to this area of research should help characterize chemical processes inside gas-phase chemical reactors.

Recent developments in Raman spectroscopy have already moved the technique to the forefront of combustion research, and there remains considerable scope for its development in, for example, the area of gaseous chemical reactions for both homogeneous and gas-phase measurements above catalytic surfaces. Conventional Raman reactions will continue to contribute fundamental data from controlled laboratory systems; but the new generation of Raman techniques, in particular CARS, will make possible the difficult task of measuring temporally and spatially precise temperatures and concentrations in practical combustors.

1. Raman, C.V. *Ind. J. Phys.* **2**, 387 (1928).
2. Régnier, P.R. & Taran, J.-P.E. *Appl. Phys. Lett.* **23**, 240 (1973).
3. Maker, P.D. & Terhume, R.W. *Phys. Rev.* **137**, A 801 (1965).
4. Blattie, J.R., Black, I.D., Gilson, T.R. & Greenhalgh, D.A. *Coherent Anti-Stoke Raman Spectroscopy in Sooty Diffusion Flames* SPIE Pap. 158-09 (1978).
5. Stenhouse, I.A., Williams, D.R., Cole, J.B. & Swords, M.D. *Appl. Optics* **18**, 3819 (1979).
6. Attal, B., Péalat, M. & Taran, J.-P.E. *CARS Diagnostics of Combustion*, presented at the AIAA 18th Aerospace Sciences Meeting, Pasadena, January 1980.
7. Eckbreth, A.C. *Comb. & Flame* **39**, 133 (1980).
8. Greenhalgh, D.A., England, W.A. & Porter, F. to be published.
9. Lapp, M. & So, R.M.C. *The Study of Turbulent Diffusion Flames: Modelling Needs and Experimental Light Scattering Capabilities*, Proc. AGARD Conf. on Testing and Measurement Techniques in Heat Transfer and Combustion, Propulsion and Energetics Panel, Brussels, May 5–7, 19–1 (1980).
10. Bechtel, J.H. & Blint, R.J. *Appl. Phys. Lett.* **37**, 576 (1980).

Amazonian question

from D.H. Tarling

THE SOUTH AMERICAN SHIELD covers a vast area of some 8,000,000 km^2, stretching from northern Venezuela to northern Argentina. To the west it is bounded by the Andes and to the east, it originally linked with Africa. However, the area is still largely unknown because of the difficulty of access and the general lack of exposure of basement rocks. The shield poses many questions of both local and international importance, for example the apparently unique persistence of Archaean-type tectonics well into Proterozoic times, which were recently discussed at the First Amazonian Symposium held at Puerto Ayacucho in Venezuela.

One of the fundamental economic problems is the apparent lack of the mineralization normally associated with Archaean terrains. The question is whether such mineralization never occurred, occurred and has not been located, or occurred but has since been eroded. In the case of the ultrabasics and greenstones in Brazil, Guyana and Surinam, the major problem appears to be that the techniques of exploration have not yet been sufficiently modified for local conditions — the great thickness (often 100 m) of laterite effectively prevents the use of most geophysical techniques while geochemical prospecting really requires a return to first principles, rather than the adoption of techniques developed for use in the medium latitudes of the Northern Hemisphere. G.W. Walrond (Guyana) identified seven distinct metallogenic provinces but the trace mineral associations that needed to be mapped may not be those used for the Northern Hemisphere where the degree and nature of hypergene weathering are so different. Nonetheless, it was clear that in this area, at least, the mineralization was present, but the current techniques were not really suited for its optimum location.

Such an assessment applies also to the use of remote sensing in such areas. In general, the vegetation in equatorial rain forests exists independently of the basement rock. It does not, therefore, show the usual stress features associated with, for example, copper impurities in the subsoil. Ecological provinces can be recognized from satellite imagery, but they are too large to be used directly in the search for minerals. Structurally controlled lineations can, however, be determined, particularly from radar scans of the region. The fact that the vegetation is self-dependent also highlights the problem of the extent of forest burning in many parts of the Amazon Basin. Once destroyed, the soils and humus are rapidly removed and the area becomes sterile. In such remote areas, the problem is not one of legislation but one of enforcement.

Elsewhere Archaean mineralization may have existed, but is now eroded. Many of the richest mineral deposits associated with Archaean terrains are, in fact, sedimentary concentrations that have since metamorphosed. The original igneous sources of copper and gold, for example, are usually in the greenstones and ultrabasics, but these are marginally economic, and it is surficial concentration that caused the real enrichment of areas such as the Copperbelt of Central Africa. Such superficial deposits may have been removed from the Venezuelan part of the Archaean — the Imataca Province — which was possibly eroded to greater depths than similar greenstone terrains elsewhere. There is one simple test of this hypothesis. Pollack and Chapman (*Tectonophysics* **38**; 279, 1977) demonstrated an age relationship between the rate of heat flow and the age of the last orogenic event in continental areas. Such a relationship requires a strong concentration of radiogenic heat-producing elements in the upper two or three kilometres of the continental crust. If the depth of erosion in northern Venezueala has been that deep, such radiogenic heat sources would have been removed and this area would now be characterized by even lower heat flow.

The most intriguing problem is the source of the diamonds that can be found extensively over the shield. These diamonds are mostly of industrial grade, but the few gemstones are of exceptional purity. It was not surprising, therefore, that Baptista and Parra's (Venezuela) report of diamond finds in the sediments of the Quebrada Grande river at 144,000 carats per km^2 caused considerable excitement. In South Africa, such concentrations would normally be very close to their kimberlite source. However, in South America these alluvial diamonds are not found associated with the other minerals usually found near kimberlite sources. Asbestos was found but is attributable to ultrabasic intrusives nearby; none of the chrome spinels associated with kimberlites was reported. It may well be asked why it is so essential to find the source as, in general, diamonds within kimberlite pipes are very expensive to extract. The reason is that knowledge of the source is required to assess areas in which other placed deposits are likely to occur. The search for

D.H. Tarling is a Reader in Palaeomagnetism, in the Department of Geophysics and Planetary Physics, Newcastle University, UK.

0028-0836/81/250537-02$01.00

Opportunism of plant seeds

from Peter D. Moore

THE germination of many types of seed can be triggered by various external, physical stimuli, such as temperature, water availability and light. Logically it would seem that such response mechanisms have developed as a consequence of a selective advantage gained by the plants whose development has previously been constrained in this way. Dormancy thus prevents the plant from germinating and expending its reserve at an inopportune moment, that is, one in which its chances of long-term survival are small.

The germination strategies of plants are therefore closely linked to the ecological problems presented by their particular habitat. Many weed species occupying unpredictable habitats, in which frequent catastrophe is an inevitable component of life, rely strongly upon the maintenance of a reservoir of dormant seeds in the soil to ensure survival of the holocaust. In such species a simple technique for dormancy maintenance is the linkage of dormancy breakage with exposure to light. Species which grow in dense grassland habitats would be at a great disadvantage if dormancy were maintained at low light intensity, for at ground level they may rarely experience anything else. The seeds of such species usually germinate in the dark. Species of woodland clearings depend on the death of a canopy tree or some other disturbance to open the leaf cover and provide an opportunity for growth. It might be supposed that such species would be at an advantage if their germination were finely tuned to the changes in spectral quality of light associated with a reduction in canopy. This is speculation, but it would make sound ecological and evolutionary sense.

Cresswell and Grime (see p.583 of this issue of *Nature*) point out an interesting correlation between the light demands of seeds and the environment within which they mature in the fruit. Seeds developing within tissues which are photosynthetic are shielded by these green layers and subsequently require a light stimulus for germination. Seeds which are exposed to full light before their final maturation and drying are capable of germinating without light.

Cresswell and Grime propose that species of unpredictable habitats which gain advantage by maintaining a dormant population of seeds in the soil, retain chlorophyllous tissues around their maturing seeds, thus producing a light requirement. It is even possible that some species whose seeds do not become incorporated in the soil seed bank may retain a green cover over maturing seed and thus induce a light requirement; such species may require the full light of an opening in the canopy to trigger germination.

Obviously, light is just one of many factors in the environment which influence germination and it is not always the overriding one. However, the experiments of Cresswell and Grime do suggest that the morphology and phenology of the inflorescence, flower and fruit can have a direct influence upon the subsequent germination requirements of the seed. This leads to the interesting possibility that ecotypic development within certain species, providing adaptation to a variety of habitats, could relate to such a minor modification as the chlorophyll content of structures investing the seed. It will be of value to look for such modifications among those plants which are capable of survival in both stable and unstable environments.

Peter D. Moore is in the Department of Plant Sciences, University of London King's College.

---

kimberlites, however, persists. Elsewhere in the world the formation of kimberlites seems to require two factors — an Archaean-type upper mantle and a later addition of volatiles to it. In South America, Archaean areas are present, but they are mostly well away from sources of volatiles, such as later subduction zone activity. Both factors may be present to the north of the Sierra de la Ventana mountains, suggesting the possibility of kimberlites in southern Brazil and Uruguay, and the repetitive subduction along the Andean chain during the past billion years may also provide a volatile source, although Archaean upper mantle materials appear to be absent along this border.

Diamonds are a component of another intriguing feature of the South American Shield — the Roraima Formation — which forms the isolated mountain tops that were the locus for Conan Doyle's 'Lost World'. The basal conglomerates of this Precambrian sedimentary sequence are generally diamondiferous, and sedimentological studies (Gosh, Venezuela) show very clearly that the entire sequence came from the east. But what is the source of these sediments and where did the diamonds come from? The total original volume must have been at least 1 million km^3. The age of the basement on which they lie is different in different areas and it is generally thought that the base of the Roraima itself is strongly diachronous. There are major differences when the sediments are traced laterally, but the scale of this deposit makes such variations inevitable. The puzzle is why the similarities should persist over 1 million km^2. Sedimentological evidence clearly rules out the suggestion by Gaudett (New Hampshire) that local Venezuelan sources were of major importance. Gaudett and Olszweski demonstrated a consistent migration in the age of the formation of the basement rocks of southern Venezuela. Strontium isotopes showed that these Proterozoic granites and gneisses had little continental contamination, but their age decreased from $\sim$ 1,850 Myr in the north-east to 1,550 Myr in the south-west. They suggested that this indicated a gradual migration of subduction activity up to a final suture zone close to the Venezuelan–Colombian border. What was important about Gaudett's study was that essentially similar techniques of whole-rock and zircon analyses were used. Many problems in understanding the geological evolution of South America arise from the wide variety of techniques that have been applied, each of which often represents the closure of a different geochemical system. Palaeomagnetic results, for example, from this and other Precambrian terrains, often assume that the radiometric age, irrespective of the method, represents the age of the magnetization. This may be true for some K–Ar methods, but will generally not be the case for other techniques.

This radiometric dating problem is particularly important, on a world scale, in any attempt to understand the change from Archaean to Proterozoic tectonic conditions. There seems to be good evidence from virtually all other shield areas that Archaean tectonic conditions ceased between 2.5 and 2.7 billion years ago (Moorbath *Phil. Trans. R. Soc.* A **288**; 401, 1978). In northern Venezuela, the Imataca Province has an age of 3.1–3.3 billion years, but the oldest age, so far, from the Guyana Shield is 2.2–2.1 billion years (Gibbs, thesis, Harvard University, 1979) — indicating the persistence of 'Archaean' greenstone formation for some 500–600 Myr longer that in the rest of the world. Unless these dates have been significantly reset, they clearly require drastic modifications to be made to fit tectonic models that involve a world-wide change from Archaean- to Proterozoic-type tectonics over a period of some 200 Myr. Such critical problems can only be resolved by further work. It is unlikely that the next meeting (Brazil 1984) will provide all the answers, but the First Amazonian Symposium has certainly set the scene within which further work can be more effectively planned. □

0028-0836/81/250538-01$01.00

REVIEW ARTICLE

Mechanisms of slow postsynaptic potentials

H. Criss Hartzell

Department of Anatomy, Emory University School of Medicine, Atlanta, Georgia 30322, USA

Classical views of synaptic transmission have to be modified to take account of slow postsynaptic potentials (p.s.ps), which are often mediated by novel ionic and molecular mechanisms and which are sometimes evoked by neuropeptides. Understanding slow p.s.ps will provide insights into the mechanisms of biological signal transduction and long-term signalling in the nervous system.

THE key to understanding synaptic transmission, neuronal integration and signalling is to discover how neurones recognize extracellular signals and transduce the signals into electrophysiological responses. Much of what we know about the recognition and transduction processes has been gathered from studies on the skeletal neuromuscular junction[1,2,104] in which acetylcholine (ACh) released from the motor nerve terminals opens specific ionic channels in the membrane by binding to postsynaptic receptors. The response of the postsynaptic membrane to ACh is rapid[3,4] (Fig. 1*a*); the conductance change is complete within 20 ms, the voltage response is slowed by the membrane time constant. Many synapses, however, do not work as rapidly. The homeostatic and autonomic functions of the nervous system, which operate with time courses of seconds or minutes, are not well served by fast synaptic responses lasting only milliseconds. For example, regulation of heart rate requires that transmitter action extend over several cycles of the beat[5]. Long transmitter action is often accomplished by slow synaptic responses which, in the heart, last several seconds (Fig. 1*b*) and in other targets may last tens of minutes (Table 1). Slow postsynaptic potentials (p.s.ps) are defined as responses which begin with a lag period of tens of milliseconds after transmitter application and are slower than the membrane time constant. Slow p.s.ps, like fast p.s.ps, may be either excitatory (e.p.s.p.) or inhibitory (i.p.s.p.) depending on whether they increase or decrease the likelihood that the cell will reach threshold and initiate action potentials. Slow p.s.ps have attracted interest because they differ from fast p.s.ps in several fundamental ways. Slow p.s.ps: (1) are probably mediated by different molecular mechanisms from those mediating fast p.s.ps, (2) are often produced by novel ionic mechanisms which involve the closing of ionic channels or modulation of voltage-sensitive channels, and (3) are sometimes produced by non-classical transmitters, notably neuropeptides. The purpose of this review is to summarize these features of slow p.s.ps and to evaluate their possible mechanisms and roles. The interested reader is referred to several extensive reviews which have recently appeared[6–12,105].

Ionic mechanisms of slow p.s.ps

Many slow p.s.ps are mediated by conventional ionic mechanisms: a transmitter opens ionic channels and increases ionic conductance (g). For example, the slow response of Fig. 1*b* is due to an increase in the conductance of potassium (g_K)[13,14]. Weight and Votava[15] first proposed, however, that some slow

Table 1 Selected examples of slow postsynaptic potentials

Tissue	Species	Response	Δ Conductance	Latency	Duration	Transmitter	Refs
Sympathetic neurones	Mammal	Slow i.p.s.p.	↓Na, electrogenic?	100 ms	1–5 s	ACh, DA?	6–9
		Slow e.p.s.p.	↓K, ↑Na, ↑Ca(?)	100–400 ms	10–60 s	ACh	6–9
	Frog	Slow i.p.s.p.	↑K	100 ms	1–5 s	ACh	6–10, 103
		Slow e.p.s.p.	↓K, ↑Na, ↑Ca(?)	100–400 ms	10–60 s	ACh	6–10, 18
		Late-slow e.p.s.p.	↓K, ↑ ?	1–5 s	10–30 min	LHRH	29–31
Parasympathetic neurones	Mudpuppy	Slow i.p.s.p.	↑K	30–100 ms	2–20 s	ACh	21
Mesenteric neurones	Guinea pig	Slow e.p.s.p.	↑(?)		2–60 s	Sub. P?	9, 34
Myenteric neurones	Guinea pig	Slow e.p.s.p.	↓K	0.5–5 s	5–90 s	Sub. P, 5-HT	9, 33, 94
		Slow i.p.s.p.	↑K	0.5–1 s	2–40 s	Non-cholinergic	33
Submucous plexus neurones	Guinea pig	Slow i.p.s.p.	↑K			catecholamine	95
L-Ped-3	*Planorbis*	Slow i.p.s.p.	↑K	100 ms	5–10 ms	ACh	48
Cortical neurones	Cat	Slow e.p.s.p.	↓K			ACh	96
L14 motor neurones	*Aplysia*	Slow e.p.s.p.	↓K		10 s–3 min		93
Medial cells	*Aplysia*	Slow e.p.s.p.	↓K	200–500 ms	5–20 s	ACh	38, 52
Neurones	*Helix*	Slow e.p.s.p.	↓K		10 s	5-HT	39
		Slow i.p.s.p.	↓K, ↓Na		20 s	5-HT	39
Cardiac muscle	Vertebrate	Slow i.p.s.p.	↑K	100–300 ms	2–30 s	ACh	5, 21, 40–43
Smooth muscle	Vertebrate	Slow e.p.s.p.	↑K, ↑Na, ↑Ca?	100–400 ms	2–30 s	ACh	11, 44, 45
Salivary gland	Cockroach	Slow i.p.s.p.	↑K	1 s	10–20 s	DA	46
Submaxillary gland	Cat	Biphasic?	↑Na, ↑K	500 ms	2–3 min	ACh	97
Parotid gland	Cat	Slow e.p.s.p.	↑Na, ↑K	1–3 s	Minutes	Catecholamine (β)	97, 98
		Slow i.p.s.p.	?	300 ms	Seconds	ACh	19, 22, 97, 98
Pancreatic acinar cells	Cat	Hyperpolarization	↑Na, ↑K	100–300 ms	2–3 min	ACh	20, 22, 97

Abbreviations: ACh, acetylcholine; DA, dopamine; LHRH, luteinizing hormone releasing hormone; 5-HT, serotonin.

0028-0836/81/250539—06$01.00

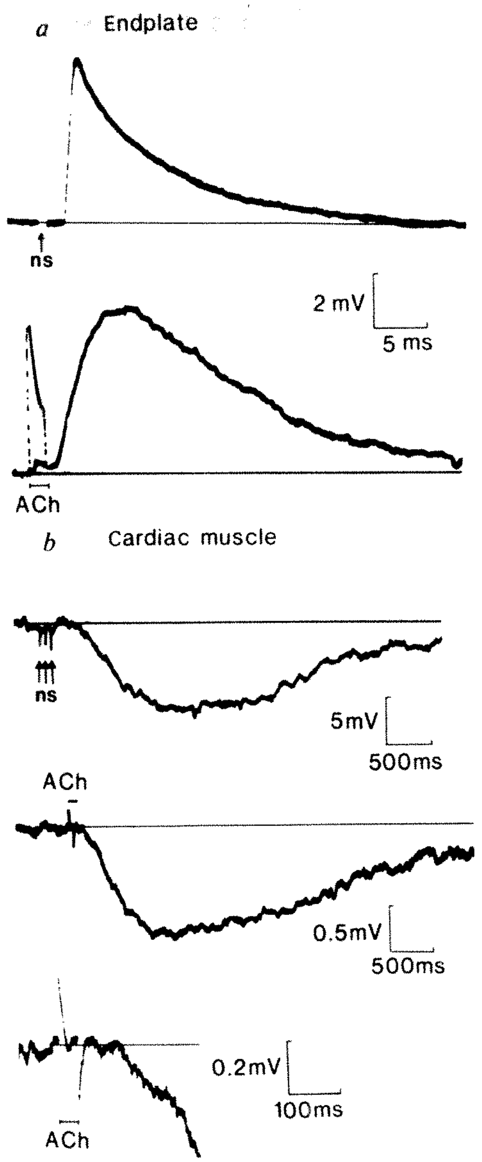


Fig. 1 Fast and slow p.s.ps. *a*, Response of frog skeletal muscle (cutaneous pectoris) to ACh. Top trace: the motor nerve was stimulated at the arrow (ns = nerve stimulation) with a 5-V, 0.5-ms duration pulse. The response of the muscle was recorded by an extracellular electrode positioned on the postsynaptic membrane. This type of recording measures the voltage drop produced by the synaptic current flowing across the resistance of the extracellular synaptic cleft. The time course of the response is virtually identical to that of the synaptic current. The time between the stimulus and the onset of the response is due to conduction time of the nerve and synaptic delay. Bottom trace: ACh was applied ionophoretically from a micropipette in a 1-ms pulse. The response is recorded with an intracellular voltage-recording electrode. The current pulse used to eject the ACh is superimposed on the intracellular record. The response begins during the 1-ms pulse of ACh. *b*, Response of muscle in sinus venosus of frog heart to ACh. Top trace: response to vagus nerve stimulation. Three shocks (arrows, ns) were delivered to the vagus nerve at 20 s^{-1} and the response recorded with an intracellular electrode. Middle trace: response to a 50-ms duration pulse of ACh released from an iontophoretic pipette. Bottom trace: same as middle trace on a faster time base and at higher amplification to illustrate latency.

p.s.ps are produced by the closing of ionic channels which are normally open in the absence of transmitter. This suggestion was based on the finding that membrane resistance, monitored by the voltage deflection produced by a pulse of current injected into the cell, increased during the depolarizing slow e.p.s.p. in bullfrog sympathetic neurones[15,16]. In addition, the reversal potential of the slow e.p.s.p. seemed to be in the direction of E_K because the amplitude of the slow e.p.s.p. usually decreased[15,16] or reversed polarity[15,17] with hyperpolarization. It was reasoned that the slow e.p.s.p. was produced by a decrease in g_K, as follows. At membrane potentials positive to E_K, K^+ ions diffuse down their concentration gradient out of the cell. When transmitter turns off this resting outward current, the cell depolarizes towards the equilibrium potentials of the other permeant ions, mainly sodium.

Recent voltage-clamp studies have clearly demonstrated that the slow e.p.s.p.[18] and the late-slow e.p.s.p.[19] in bullfrog sympathetic ganglia involve turn-off of a unique K^+ current ('M-current'). An interesting feature of the M-current is that the channels are sensitive to membrane potential as well as to transmitter[18]. Near resting potential (−30 to −60 mV) the M-current channels are open, but they are closed at membrane potentials negative to −60 mV. In contrast, the conductance of channels responsible for fast p.s.ps, such as ACh-activated channels at the motor endplate, are not affected by membrane potential, although channel open-time is somewhat affected by voltage[20].

Voltage sensitivity is not restricted to p.s.ps produced by conductance decreases and may be a common property of slow p.s.ps. For example, a muscarinic slow i.p.s.p. produced by increased g_K in mudpuppy parasympathetic neurones is voltage sensitive[21]. In the range of −100 mV to −60 mV, the ACh-induced increase in g_K is constant, but thereafter it decreases, until above −40 mV the channels are inactivated. This suggests that excitatory transmitters, including ACh acting on nicotinic receptors on the same neurone, have the power to inactivate and overide the muscarinic slow i.p.s.p. With both the slow e.p.s.p. (above) and the slow i.p.s.p, the conductance change produced by transmitter decreases as the membrane potential moves away from the reversal potential of the response (that is, as the driving force of the response increases). Thus, transmitter will be most effective at a membrane potential between the reversal potential of the response and the membrane potential where channels are inactivated.

The ionic mechanisms of many slow p.s.ps have been difficult to elucidate. For example, because M-channels are inactivated at membrane potentials negative to −60 mV, accurate determination of the reversal potential of the slow e.p.s.p. has been difficult. Most investigators find that the amplitude of the slow e.p.s.p. decreases with hyperpolarization[15,17,23,24], as expected for a pure g_K decrease with reversal potential near E_K. However, a few investigators report that hyperpolarization increases the amplitude of the slow e.p.s.p. and that the slow e.p.s.p. is accompanied by increases in membrane conductance in some cells[23–26]. To explain these discrepancies, Kuba and Koketsu[6] have proposed that the slow e.p.s.p. is produced both by a decrease in g_K and by increases in conductance of Na^+ and Ca^{2+}. When the cell is hyperpolarized below −60 mV, conductance increases are more noticeable because M-current is inactivated and driving forces on Na^+ and Ca^{2+} conductances increase. The ratio of the decrease and increases in conductance apparently varies in different cells[6]. Possibly, discrete populations of cells exist with different ionic mechanisms subject to different control, mediating their slow e.p.s.ps. Likewise, the ionic mechanisms of both the late-slow e.p.s.p. and the slow i.p.s.p. in sympathetic ganglia are confusing[9,27,28,103].

Peptides as neurotransmitters producing slow p.s.ps

The neuroactive peptides in peripheral nerves and ganglia may be responsible for slow p.s.ps. Good evidence for this hypothesis comes from Jan *et al.*[27,29,30] working on the role of lutenizing hormone releasing hormone (LHRH) in the late-slow e.p.s.p.[31] of bullfrog sympathetic neurones. Stimulation of the 8th and 9th spinal nerves in the presence of cholinergic antagonists produces a depolarization which may last 10–20 min, is caused partly by a decrease in g_K and is mimicked by the application of LHRH and certain analogues. Immunoreactive LHRH is concentrated in the preganglionic axons which synapse on the sympathetic neurones[30]. In view of this evidence and the fact that LHRH is released from the ganglion in a calcium-dependent manner when the ganglion is depolarized with high K$^+$, LHRH meets the minimal criteria for a transmitter.

Peptides may also be transmitters in other ganglia. Substance P has been identified in nerve terminals of the guinea pig myenteric plexus[32]. Application of the peptide mimics the effects of stimulation of the nerves, producing a slow, decreased-conductance e.p.s.p. in the myenteric neurones[33]. Substance P may also mediate a slow e.p.s.p. in mesenteric nerones in the guinea pig[9,34]. Other peptides have been shown to produce membrane potential changes or modulate synaptic potentials in vertebrate autonomic ganglia and in invertebrate ganglia, but little direct evidence exists to support their roles as transmitters responsible for slow p.s.ps.

Molecular mechanisms of slow p.s.ps

The sequence of events leading to neurotransmitter-induced membrane conductance changes can be divided arbitrarily into four steps: (1) diffusion of transmitter to the receptive membrane, (2) binding of transmitter to receptor, (3) opening or closing of the ionic channel, which is presumed to be due to a conformational change in the channel, and (4) the relaxation of the channel to its initial closed or open conformation. In fast synaptic responses, such as the endplate current, the entire process is completed within tens of milliseconds[3]. Diffusion of ACh from the nerve terminal to the receptors takes place in less than 50 µs, and the binding of ACh to receptors and channel opening are largely completed within 300 µs (ref. 3). The time course of decay of the endplate current is due to the first-order closing of channels, which have exponentially distributed open times with a mean equal to the time constant of endplate current decay[20,35–37]. Which of these steps is rate-limiting for slow p.s.ps? Because the time course of many nerve-evoked slow p.s.ps is mimicked by short iontophoretic pulses of transmitter[3,11,21,38–46], the slowness of these responses is almost certainly a postsynaptic phenomenon; but the possibility remains that very long-lasting p.s.p.s produced by neuropeptides may be due in part to prolonged transmitter release[105].

Diffusion. The time course of slow p.s.ps is much slower than free diffusion in bulk solution[5,21,40,42,43,46,47], but the onset of slow p.s.ps might be slowed by diffusion barriers restricting access of transmitter to receptors[11,12,40,41,47]. A variety of experiments argue against such a mechanism. (1) At several slow synapses no morphological diffusion barrier has been identified. For example, in cardiac muscle, where ACh produces a slow i.p.s.p. (Fig. 1*b*), receptors are diffusely distributed over the entire surface of the muscle and are not localized to clefts or other regions which appear inaccessible to transmitter[41]. (2) The same transmitter can produce both fast and slow p.s.p.s in the same neurone. In parasympathetic neurones of the mudpuppy[21], nicotinic ACh receptors mediate a fast e.p.s.p. which is complete in 100 ms, whereas muscarinic ACh receptors mediate a slow i.p.s.p. which is delayed 100 ms or more and lasts several seconds. Both nicotinic and muscarinic receptors are located on the neuronal soma. It seems unlikely that diffusion of ACh to muscarinic receptors is 100–1,000 times slower than its diffusion to nicotinic receptors. (3) The onset and latencies of slow p.s.ps are much more temperature sensitive than would be expected if diffusion were rate-limiting[12,21,40,41,46,48,49]. Temperature coefficients (Q_{10}) as high as 10–12 have been reported. (4) Diffusion barriers should be overcome by large doses of transmitter. Latencies of slow p.s.ps, however, are rather insensitive to transmitter dose[40,41]. (5) Nargeot *et al.*[50] have recently developed a technique to activate receptors in a way which virtually eliminates diffusion time of agonist to receptor. Tissue is equilibrated in agonist and a photosensitive antagonist. The antagonist can then be rapidly inactivated (<1 ms) by a flash of light. This exposes previously blocked receptors to agonist, which has been equilibrated with the tissue. In atrial muscle the increase in the ACh-induced K$^+$ current, caused by rapid inactivation of the antagonist *trans*-3-3′-*bis*-(α-(trimethylammonium)methyl)azobenzene in the presence of ACh, has the same time course as the response produced by iontophoretic application of ACh.

Binding of transmitter to receptor. Hill-Smith and Purves[40] have calculated that the kinetics of transmitter binding can explain the decay of the slow p.s.p. In tissues where receptors are diffusely distributed[41] the slow decay of p.s.ps might easily be explained by diffusion of transmitter along the receptive membrane and repeated binding to receptors. The quantitative effect of binding on the time course of the p.s.p., however, depends on the on- and off-rates of transmitter–receptor interaction. Kinetics of labelled agonist binding has not been published, but dissociation of muscarinic agonists is slow enough for equilibrium binding to be measured in an assay[51] in which membranes are washed several times with buffer to remove free radiolabelled ligand, a process which requires several seconds. This suggests that kinetics of dissociation of agonist could explain the decay of slow p.s.ps.

This hypothesis is supported by the finding that the time course of the muscarinic hyperpolarization in cardiac muscle reflects the affinity of agonist binding to receptor. For example, Birdsall *et al.*[51] have shown that oxotremorine binds to low- and high- affinity sites in brain with 10- and 21-fold higher affinity, respectively, than does carbamylcholine (CCh). In frog cardiac muscle, oxotremorine is 66 times more potent in competing for ^{3}H-QNB (quinuclidinyl benzilate) binding than is CCh[41]. The hyperpolarization produced by iontophoretic application of oxotremorine is considerably slower (half decay time, $t_{\frac{1}{2}} = 5.1$ s) than the hyperpolarization produced by CCh ($t_{\frac{1}{2}} = 1.2$ s) (Fig. 2). The finding that oxotremorine has a higher affinity for receptor and produces a slower hyperpolarization favours the idea that the kinetics of agonist dissociation is partly responsible for the slowness of this p.s.p.

Although the decay of some slow p.s.ps may be accounted for by slow transmitter dissociation, the latency of slow p.s.ps is not easily explained[40] by the kinetics of agonist binding unless several steps occur in series between agonist binding and channel opening[12,21,52]. These steps could involve the production of second messengers or cooperative interaction of several receptors in the opening of a channel.

Activation of conductance changes. The channels responsible for some slow p.s.ps may have very long open-times. Marty and Ascher[52] have shown that an ACh-induced synaptic current in an *Aplysia* neurone relaxes very slowly in response to voltage jumps and have concluded that the mean channel open-time is several seconds.

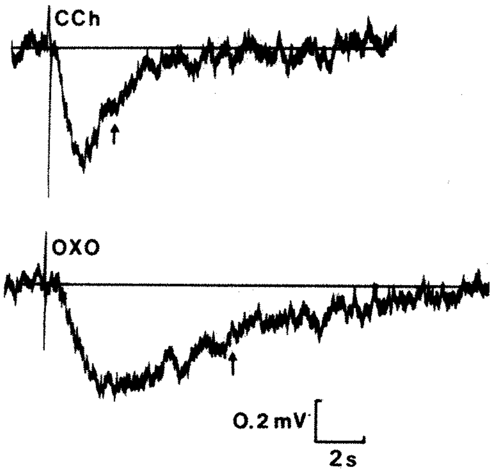


Fig. 2 Time course of responses of frog sinus venosus to muscarinic agonists. Upper trace: response to a 50-ms duration iontophoretic pulse of carbamylcholine (CCh). Bottom trace: response to a 5-ms duration iontophoretic pulse of oxotremorine (oxo). The pipette contained 0.5 M oxotremorine neutralized to *p*H 8 with HCl. Responses to both CCh and oxotremorine were on the linear portion of the dose–response curve. This was important because the time course of responses to muscarinic agonists applied iontophoretically is dose dependent[46]. Arrows indicate time at which response declines to half its peak amplitude. Time from peak to half-decline time is $t_{\frac{1}{2}}$.

On the other hand, a mean channel open-time of $\sim$100 ms for the ACh-activated K^+ channels in rabbit sinoatrial node has been inferred from analysis of ACh-induced noise[53,54] and from relaxation of the ACh-induced current[55]. Channel open-times vary somewhat with temperature, membrane potential and ACh concentration[54]. Because mean channel open-time is short relative to the decay (several seconds) of the iontophoretic response, the first-order closing of channels with lifetimes of $\sim$100 ms cannot be rate-limiting and some channels must be activated during the decay of the p.s.p. Repeated channel opening could be due to slow dissociation of agonist from the receptor or repeated agonist binding.

Channels which have been studied by single-channel recording usually have two states, open and closed, and transitions between these states are rapid[37]. The onset of the slow p.s.p., therefore, probably reflects a slow increase in the fraction of open channels after transmitter binding rather than a slow transition of individual channels from one state to another. If transmitter binding instantaneously alters the opening (β) and closing (α) rate constants of the channel, the fraction of open channels will change exponentially with a time constant of $(\alpha'+\beta')^{-1}$, where α' and β' are the new rate constants. The estimates for α ($10.1\ s^{-1}$) and β ($12.0\ s^{-1}$) for the ACh-induced K^+ current in rabbit sinoatrial node[54] predict a time constant of $\sim$45 ms, which is significantly less than what is observed for the onset of the response. Thus, factors other than channel kinetics probably contribute to the slowness of the onset (but compare ref. 56). The slow onset could be explained if channel opening is not mediated by direct transmitter-induced conformational changes of the channel but is caused indirectly by second messengers which are generated by agonist binding and build up slowly in the region of the channel.

It has been suggested that the second messengers mediating slow p.s.ps are cyclic nucleotides[57]. In this scheme, transmitter activates an adenylate cyclase, which catalyses the production of cyclic AMP; cyclic AMP then activates a cyclic AMP-dependent protein kinase which phosphorylates the ionic channel. The channel is postulated to be open when phosphorylated and closed when dephosphorylated, or vice versa. An analogous mechanism may involve cyclic GMP[57]. Good evidence exists for the involvement of cyclic AMP in mediating the increases in the action potential Ca^{2+} current produced by β-adrenergic agonists in cardiac muscle (reviewed in refs 12, 58) and by serotonin in an *Aplysia* neurone[59–67]. In both examples, (1) transmitter causes an increase in cyclic AMP production in the ganglia, (2) extracellular application of cyclic AMP derivatives or intracellular injection of cyclic AMP causes an increase in the Ca^{2+} current, and (3) agents, such as phosphodiesterase inhibitors, which increase cyclic AMP levels in the target, potentiate the electrophysiological effect of the transmitter. In the *Aplysia* neurone the electrophysiological action of serotonin correlates with increased ^{32}P incorporation into a 137,000 molecular weight membrane protein[61]. Further, intracellular injection of the purified catalytic subunit of cyclic AMP-dependent protein kinase mimics the electrophysiological effects of serotonin[60].

At other synapses the role of cyclic nucleotides is unresolved. For example, the concept that cyclic AMP mediates slow p.s.ps was advanced largely by the work of Greengard[57] and co-workers on the slow i.p.s.p. in neurones of sympathetic ganglia. Although impressive evidence has been amassed in support of this idea, this hypothesis has recently been challenged by several laboratories. The evidence has been evaluated by others[9,10,28,62,63] and is summarized in Table 2. Perhaps the most serious problem is that only one laboratory[64,65] of several which have tried[10,63,66–74] has succeeded in mimicking the hyperpolarizing effect of the presumed transmitter, dopamine, by extracellular or intracellular application of cyclic AMP or cyclic AMP analogues. Furthermore, in the positive experiments[64,65] no evidence was presented to show that the hyperpolarizations produced by cyclic AMP and dopamine occurred by identical ionic mechanisms. Another potential problem is that the transmitter which produces the slow i.p.s.p. may not be dopamine, but ACh[9,28,75,76,103]. If the natural transmitter is not dopamine, the basic assumption of many of the tests has been invalidated. Another controversy concerns whether the slow e.p.s.p. in sympathetic neurones is mediated by cyclic GMP[77–80]. Although evidence suggests that cyclic AMP may mediate membrane changes at some synapses, proof of this hypothesis must await further tests, including demonstration that transmitter-induced changes in membrane conductance are accompanied by cyclic AMP-dependent phosphorylation of appropriate membrane proteins.

Speculation about other second messengers

Second messengers are attractive mechanisms for explaining certain features of slow p.s.ps described in this review. As cyclic nucleotides may be involved in only certain p.s.ps, I will consider several other second messenger systems which could possibly have a role in mediating slow p.s.ps.

Information has accumulated that certain neurotransmitters, including those acting on α-adrenergic and muscarinic receptors, and some neuropeptides stimulate phosphatidylinositol breakdown to diacylglycerol and phosphatidic acid[81]; this has gained possible significance by the finding that phosphatidic acid can function as a calcium ionophore in artificial membranes[82]. Thus, phosphatidic acid produced by phosphatidylinositol breakdown in response to transmitter might act as a calcium pore in cell membranes. This suggestion is supported by the finding that exogenously applied phosphatidic acid produces contraction in smooth muscle cells[83]. Further, the ACh-induced increase in the Ca^{2+}-activated K^+ permeability of parotid gland cells is mimicked by application of phosphatidic acid and calcium[84]. In neutrophils, however, phosphatidylinositol breakdown seems to be a consequence and not a cause of Ca^{2+} influx[85].

The production of diacylglycerol as a result of phosphatidylinositol breakdown may also act as a second messenger in activating a novel protein kinase ('C-kinase') recently discovered by Nishizuka and co-workers[80]. C-kinase could function to alter membrane permeability in an analogous way to that proposed for cyclic nucleotide-dependent protein kinases. It requires calcium and phosphatidylserine for its enzymatic activity[86]. Enzyme activity is markedly stimulated by diacylglycerol[87,88], which increases the affinity of C-kinase for Ca^{2+} about 20-fold. Nishizuka has proposed that certain neurotransmitters activate membrane phospholipase C, which catalyses phosphatidylinositol breakdown to diacylglycerol. Diacylglycerol activates C-kinase in the presence of the low

Table 2 Mediation of the slow i.p.s.p. by cyclic AMP

Evidence in favour	Selected refs
Preganglionic nerve stimulation or dopamine application causes an increase in cyclic AMP levels in ganglion	64, 99–101
Increase in cyclic AMP is localized to ganglion cell soma	102
Dopamine-sensitive adenylate cyclase is present in ganglion	100
Methylxanthine phosphodiesterase inhibitors potentiate slow i.p.s.p	63–66, 101
Extracellular application of cyclic AMP producess a hyperpolarization	64, 65
Evidence against	
Dopamine is less effective than preganglionic nerve stimulation or cholinergic agonists in producing increases in cyclic AMP	67, 101
Phosphodiesterase inhibitors which are not methylxanthines do not potentiate the slow i.p.s.p	10, 63, 67
Methylxanthines have electrophysiological effects which differ from the effects of dopamine or nerve stimulation	10, 63, 66, 68
Extracellular application of cyclic AMP does not produce a hyperpolarization	10, 63, 66–72
Intracellular injection of cyclic AMP does not produce a hyperpolarization	10, 68, 71, 73, 74
Transmitter which mediates the slow i.p.s.p. may not be dopamine	9, 28, 75, 76, 103

intracellular concentrations of Ca^{2+}. No direct evidence links C-kinase with slow p.s.ps; however, the relationship of this enzyme to phospholipid turnover, which has been puzzling for many years, and the very high concentration of C-kinase in brain make the involvement of this enzyme an attractive possibility.

Phospholipids may also effect slow p.s.ps in another way. Axelrod and colleagues[89] have demonstrated that activation of β-adrenergic receptors in several systems can stimulate methyltransferases which catalyse methylation of phosphatidylethanolamine to phosphatidylcholine in the membrane. Phospholipid methylation may mediate ligand-induced changes in membrane permeability[90] (but see ref. 91). In mast cells, IgE binding to specific cell-surface receptors stimulates phospholipid methylation, calcium influx and histamine release[90]. Inhibitors of the methyltransferases reduce phospholipid methylation by 90% and calcium influx by 85% (ref. 90). The mechanism by which phospholipid methylation affects calcium permeability is not known, although changes in membrane fluidity seem to be involved. Whether such a mechanism actually has a role in slow p.s.ps remains to be discovered.

Functional consequences of slow p.s.ps

The process of neuronal integration involves spatial and temporal summation of p.s.ps impinging on the neurone. Temporal summation of fast p.s.ps which last milliseconds requires a high degree of synchrony of converging inputs. Therefore perhaps the most important and obvious function of slow p.s.ps is to increase the length of time a synapse is active and to maximize the likelihood of temporal interaction of responses initiated at different times[5]. In addition, decreased conductance p.s.ps are particularly effective in modulating the efficacy of other inputs. For example, in sympathetic ganglia the ability of fast e.p.s.ps to initiate action potentials is increased during the late-slow e.p.s.p.[27,92] due to (1) increased membrane resistance resulting in a greater voltage deflection for the same synaptic current provided by the fast e.p.s.p., (2) summation of fast and late-slow e.p.s.ps to produce a greater depolarization, and (3) a lowered threshold for action potential generation. Activation of the late-slow e.p.s.p. can modulate the efficacy of fast e.p.s.ps from another input for up to 30 min. In addition, decreased conductance p.s.ps increase the membrane space constant and thus increase the area of the cell over which synaptic responses can interact.

The idea that such long-lasting heterosynaptic modulation has a role in behavioural responses has been demonstrated by Carew and Kandel[93]. A decreased conductance e.p.s.p. in *Aplysia* increases the effectiveness of a previously ineffective stimulus in initiating inking behaviour. Noxious stimulation of the head of *Aplysia* produces both a fast e.p.s.p. and a slow, decreased-conductance e.p.s.p. in the motor neurones which evoke inking. Inking is usually triggered by the fast e.p.s.ps. The slow e.p.s.p. increases the likelihood that a stimulus to the siphon, which normally produces subthreshold e.p.s.ps in the motor neurone, will produce inking if siphon stimulation occurs during the slow e.p.s.p. In addition, the slow e.p.s.p. increases spatial summation of fast p.s.ps by increasing membrane resistance and space constant and the electrical coupling between motor neurones.

The abilities of neurotransmitters to regulate channels which are also gated by voltage and to modulate the effectiveness and summation of other synaptic events are important features associated with slow p.s.ps. Understanding these features will surely provide insight into mechanisms of neuronal integration and synaptic plasticity.

Thanks to Joanna Webb, Vicki Shadix, Jan Zwemer-Collins, Doug Carr, Russell Fricke, Stephen Schneider, John Wood and John Scott for help with and comments on the manuscript, and to Fang for encouragement. Research in my laboratory is funded by NIH Research Career Development Award HL00435 and research grants HL21195 and HL25090.

1. Colquhoun, D. *A. Rev. Pharmac.* **15,** 307–325 (1975).
2. Steinbach, J. H. & Stevens, C. F. in *Handbook of Frog Neurobiology* (eds Llinas, R. Precht, W.) 33–92 (Springer, Berlin, 1976).
3. Katz, B. *The Release of Neural Transmitter Substances* (Thomas, Springfield, 1969).
4. Kuffler, S. W. & Yoshikami, D. *J. Physiol., Lond.* **244,** 703–730 (1975).
5. Purves, R. D. *Nature* **261,** 149–151 (1976).
6. Kuba, K. & Koketsu, K. *Prog. Neurobiol.* **11,** 77–169 (1978).
7. Nishi, S. in *The Peripheral Nervous System* (ed. Hubbard, J. I.) 225–255 (Plenum, New York, 1974).
8. Libet, B. *Fedn Proc.* **29,** 1945–1956 (1970).
9. Dun, N. J. *Fedn Proc.* **39,** 2982–2989 (1980).
10. Weight, F. F., Schulman, J. A., Smith, D. A. & Busis, N. A. *Fedn Proc.* **38,** 2084–2094 (1979).
11. Bolton, T. D. *Physiol. Rev.* **59,** 606–718 (1979).
12. Kehoe, J. S. & Marty, A. *A. Rev. Biophys. Bioengng* **9,** 437–465 (1980).
13. Hutter, O. F. in *Nervous Inhibition* (ed. Florey, E.) 144–123 (Pergamon, New York, 1961).
14. Garnier, D., Nargeot, J., Ojeda, C. & Rougier, O. *J. Physiol., Lond.* **274,** 381–396 (1978).
15. Weight, F. F. & Votava, J. *Science* **170,** 755–758 (1970).
16. Kobayashi, H. & Libet, B. *J. Physiol., Lond.* **208,** 353–372 (1970).
17. Weight, F. F. & Votava, J. *Science* **172,** 503–504 (1971).
18. Brown, D. A. & Adams, P. R. *Nature* **283,** 673–675 (1980).
19. Adams, P. R. & Brown, D. A. *Br. J. Pharmac.* **68,** 353–355 (1980).
20. Anderson, C. R. & Stevens, C. F. *J. Physiol., Lond.* **235,** 655–691 (1973).
21. Hartzell, H. C., Kuffler, S. W., Stickgold, R. & Yoshikami, D. *J. Physiol., Lond.* **271,** 817–846 (1977).
22. Kobayashi, H. & Libet, B. *Proc. natn. Acad. Sci. U.S.A.* **60,** 1304–1311 (1968).
23. Kuba, K. & Koketeu, K. *Jap. J. Physiol.* **26,** 651–669 (1976).
24. Nishi, S., Soeda, H. & Koketsu, *Life Sci.* **8,** 33–42 (1969).
25. Libet, B. *Science* **172,** 503–504 (1971).
26. Kobayashi, H. & Libet, B. *Life Sci.* **14,** 1871–1883 (1974).
27. Jan, Y. N., Jan, L. Y. & Kuffler, S. W. *Proc. natn. Acad. Sci. U.S.A.* **77,** 5008–5012 (1980).
28. Gallagher, J. P., Shinnick-Gallagher, P., Cole, A. E., Griffith III, W. H. & Williams, B. J. *Fedn Proc.* **39,** 3009–3015 (1980).
29. Jan. Y. N., Jan, L. Y. & Kuffler, S. W. *Proc. natn. Acad. Sci. U.S.A.* **76,** 1501–1505 (1979).
30. Jan, L. Y., Jan, Y. N. & Brownfield, M. S. *Nature* **288,** 380–382 (1980).
31. Nishi, S. & Koketsu, K. *J. Neurophysiol.* **31,** 109–121 (1968).
32. Pearse, A. G. & Polark, J. M. *Histochemistry* **41,** 373–375 (1975).
33. Johnson, S. M., Katayama, Y. & North, R. A. *J. Physiol., Lond.* **301,** 505–516 (1980).
34. Dun, N. J. & Karczmar, A. G. *Neuropharmacology* **18,** 215–218 (1979).
35. Katz, B. & Miledi, R. *J. Physiol., Lond.* **224,** 665–699 (1972).
36. Neher, E. & Sakmann, B. *Nature* **260,** 799–802 (1976).
37. Neher, E. in *Techniques in Cellular Physiology* (ed. Baker, P. F.) (Elsevier, Amsterdam, 1980).
38. Kehoe, J. *J. Physiol., Lond.* **225,** 147–172 (1972).
39. Gerschenfeld, H. M. & Paupardin-Tritsch, D. *J. Physiol., Lond.* **243,** 457–481 (1974).
40. Hill-Smith, I. & Purves, R. D. *J. Physiol., Lond.* **279,** 31–54 (1978).
41. Hartzell, H. C. *J. Cell Biol.* **86,** 6–20 (1980).
42. Glitsch, H. G. & Pott, L. *J. Physiol., Lond.* **279,** 655–668 (1978).
43. Pott, L. *Pflügers Arch. ges. Physiol.* **380,** 71–77 (1979).
44. Bolton, T. B. *Br. med. Bull.* **35,** 275–283 (1979).
45. Bennett, M. R. *Autonomic Neuromuscular Transmission* (Cambridge University Press, 1972).
46. Blackman, J. G., Ginsborg, B. L. & House, C. R. *J. Physiol., Lond.* **287,** 67–80 (1979).
47. Purves, R. D. *J. theor. Biol.* **65,** 327–344 (1977).
48. Katchman, A. N., Ger, B. A. & Zeimal, E. V. *Gen. Pharmac.* **11,** 55–64 (1980).
49. Creed, K. E. & McDonald, I. R. *J. Physiol., Lond.* **247,** 521–535 (1975).
50. Nargeot, J. *et al. Biophys. J.* **33,** 15a (1981).
51. Birdsall, N. J. M., Burgen, A. S. V. & Hulme, E. C. *Molec. Pharmac.* **14,** 723–736 (1978).
52. Marty, A. & Ascher, P. *Nature* **274,** 494–497 (1978).
53. Noma, A., Peper, K. & Trautwein, W. *Pflügers Arch. ges. Physiol.* **381,** 225–262 (1979).
54. Osterrieder, W., Noma, A. & Trautwein, W. *Pflügers Arch ges. Physiol.* **386,** 101–109 (1980).
55. Noma, A. & Trautwein, W. *Pflügers Arch. ges. Physiol.* **377,** 193–200 (1978).
56. Osterrieder, W., Yang, Q.-F. & Trautwein, W. *Pflügers Arch. ges. Physiol.* (in the press).
57. Greengard, P. *Science* **199,** 146–152 (1978).
58. Tsien, R. W. *Adv. Cyclic Nucleotide Res.* **8,** 363–420 (1977).
59. Klein, M., Shapiro, E. & Kandel, E. R. *J. exp. Biol.* **89,** 117–157 (1980).
60. Castellucci, V. F. *et al. Proc. natn. Acad. Sci. U.S.A.* **77,** 7492–7496 (1980).
61. Paris, C. G., Castellucci, V. F., Kandel, E. R. & Schwartz, J. H. in *Cold Spring Harb. Conf. Cell Proliferation* **8,** in press (1981).
62. McAfee, D. A. *et al. Fedn Proc.* **39,** 2997–3002 (1980).
63. Libet, B. *Life Sci.* **24,** 1043–1058 (1979).
64. McAfee, D. A. in *SIF Cells* (ed. Eranko, O.) 132–142 (U.S. Govt Printing Office, 1975).
65. McAfee, D. A. & Greengard, P. *Science* **178,** 310–312 (1972).
66. Dun, N. J. & Karczmar, A. G. *J. Pharmac. exp. Ther.* **202,** 89–96 (1977).
67. Brown, D. A., Caulfield, M. P. & Kirby, P. J. *J. Physiol., Lond.* **290,** 441–451 (1979).
58. Busis, N. A., Weight, F. F. & Smith, P. A. *Science* **200,** 1079–1081 (1978).
69. Dun, N. J., Kaibara, K. & Karczmar, A. G. *Science* **197,** 778–779 (1977).
70. Akasu, T. & Koketsu, K. *Br. J. Pharmac.* **60,** 331–336 (1977).
71. Kobayashi, H., Hashiguchi, T. & Ushiyama, N. S. *Nature* **271,** 268–270 (1978).
72. Busis, N. A., Schulman, J. A., Smith, P. A. & Weight, F. F. *Br. J. Pharmac.* **62,** 378P–379P (1978).
73. Gallagher, J. P. & Shinnick-Gallagher, P. *Science* **198,** 851–852 (1977).
74. Weight, F. F., Smith, P. A. & Schulman, J. A. *Brain Res.***158,** 197–202 (1978).
75. Weight, F. F. & Padjen, A. *Brain Res.* **55,** 219–224 (1973).
76. Weight, F. F. & Padjen, A. *Brain Res.* **55,** 225–228 (1973).
77. Weight, F. F., Petzold, G. & Greengard,P. *Science* **186,** 942–944 (1974).
78. Hishiguchi, T., Ustiyama, N. S. & Kobayashi, H. *Nature* **271,** 267–268 (1968).
79. Dun, N. J., Kaibara, K. & Karczmar, A. G. *Neuropharmacology* **16,** 715–717 (1977).
80. Dun, N. J., Kaibara, K. & Karczmar, A. G. *Brain Res.* **150,** 658–661 (1978).
81. Michell, R. *Biochim. biophys. Acta* **415,** 81–147 (1975).
82 Tyson, C. A., Vande Zande, H. & Green, D. E. *J. biol. Chem.* **251,** 1326–1332 (1976).
83. Salmon, D. M. & Honeyman, T. W. *Nature* **284,** 344–345 (1980).
84. Putney, J. W., Weiss, S. J., Van De Wolle, C. M. & Haddas, R. A. *Nature* **284,** 345–347 (1980).
85. Cockcroft, S., Bennett, J. P. & Gomperts, B. D. *Nature* **288,** 275–277 (1980).
86. Takai, Y., Kishimoto, A., Iwasa, Y., Kawahara, Y., Mori, T. & Nishizuka, Y. *J. biol. Chem.* **254,** 3692–3695 (1979).
87. Takai, Y., Kishimoto, A., Kikkawa, U., Mori, T. & Nishizuka, Y. *Biochem. biophys. Res. Commun.* **91,** 1218–1225 (1979).

88. Kishimoto, A., Takai, Y., Mori, T., Kikkawa, U. & Nishizuka, Y. *J. biol. Chem.* **255,** 2273–2276 (1980).
89. Hirata, F. & Axelrod, J. *Science* **209,** 1082–1090 (1980).
90. Ishizaka, T., Hirata, F., Ishizaka, K. & Axelrod, J. *Proc. natn. Acad. Sci. U.S.A.* **77,** 1903–1906 (1980).
91. Vance, D. E. & de Kruijff, B. *Nature* **288,** 277–278 (1980).
92. Schulman, J. A. & Weight, F. F. *Science* **194,** 1437–1439 (1976).
93. Carew, T. J. & Kandel, E. R. *Science* **192,** 150–153 (1976).
94. Wood, J. D. & Mayer, C. J. *J. Neurophysiol.* **42,** 582–593 (1979).
95. Hirst, G. D. S. & Silinsky, E. M. *J. Physiol., Lond.* **251,** 817–832 (1975).
96. Krnjevic, K., Pumain, R. & Renaud, L. *J. Physiol., Lond.* **215,** 247–268 (1971).
97. Petersen, O. H. *Physiol. Rev.* **56,** 535–577 (1976).
98. Emmelin, N., Grampp, W. & Thesleff, P. *J. Physiol., Lond.* **302,** 183–195 (1980).
99. McAfee, D. A., Schorderet, M. & Greengard, P. *Science* **171,** 1156–1158 (1971).
100. Kebabian, J. W. & Greengard, P. *Science* **174,** 1346–1349 (1971).
101. Kalix, P., McAfee, D. A., Schorderet, M. & Greengard, P. *J. Pharmac. exp. Ther.* **188,** 676–686 (1974).
102. Kebabian, J. W., Bloom, F. E., Steiner, A. L. & Greengard, P. *Science* **190,** 157–159 (1975).
103. Dodd, J. & Horn, S. *Nature* (submitted).
104. Adams, P. R. *J. Membrane Biol.* **58,** 161–174 (1981).
105. Kuffler, S. W. *J. exp. Biol.* **81,** 257–286 (1980).

ARTICLES

Graphite–magnetite aggregates in ordinary chondritic meteorites

Edward R. D. Scott, G. Jeffrey Taylor, Alan E. Rubin, Akihiko Okada* & Klaus Keil

Institute of Meteoritics and Department of Geology, University of New Mexico, Albuquerque, New Mexico 87131, USA

Aggregates consisting largely of submicrometre graphite and magnetite crystals have been identified in 14 unequilibrated and regolith breccia, ordinary chondrites. Like other primitive components, they formed before the meteorites accreted. Chondrite clasts rich in graphite–magnetite occur in three regolith breccias and are a new kind of unequilibrated ordinary chondrite. They were probably the source of graphite–magnetite aggregates in regolith breccias.

ORDINARY chondrites are believed to have formed from three major components which can easily be identified in the most unequilibrated chondrites[1,2]: chondrules, metallic Fe,Ni, and fine-grained, Si-rich, opaque (and recrystallized) matrix—'Huss matrix'[3]. These components formed in different conditions before the accretion of the meteorites. We have discovered a new component in all three ordinary chondrite groups consisting largely of an aggregate of submicrometre sized grains of graphite and magnetite (Fe_3O_4). The composition and distribution of the graphite–magnetite component were studied by optical microscopy, electron-microprobe analysis and X-ray powder diffraction.

We have found the graphite–magnetite component in three types of occurrences: as abundant isolated inclusions in eight ordinary-chondritic, regolith breccias ('gas-rich' breccias), as the sole matrix in a new kind of unequilibrated chondrite that forms clasts in these regolith breccias, and together with Huss matrix in six unequilibrated ordinary chondrites. We suggest that this component formed by low-temperature, gas–solid reactions before the accretion of the meteorites. We also suggest that the isolated inclusions of graphite–magnetite in regolith breccias were derived from bodies composed of the new kind of chondrite that has graphite–magnetite as its sole matrix.

Experimental findings

Twenty isolated graphite–magnetite inclusions were found in the Dimmitt H chondrite; they range in diameter from 30 μm to 1 mm (Fig. 1*a*). Most graphite crystals are submicrometre in size, but some are as large as 5 μm and show the characteristic optical anisotropy, bireflectance and softness of graphite. X-ray diffraction of a sample scratched from a millimetre-sized inclusion (Table 1) shows lines from graphite and a spinel, which was identified as magnetite by electron-probe analysis (Table 2). Table 1 also shows lines of aragonite ($CaCO_3$), but because $CaCO_3$ occurs in corrosion veins in the inclusions and we found low-Ca concentrations between these veins (Table 2), we believe the aragonite is a terrestrial contaminant (Dimmitt is a meteorite find).

Accurate electron-probe analyses of these inclusions were not possible because the crystals are smaller than the electron beam width (1–2 μm), and because of great uncertainties in the correction procedures for carbon analyses[4]. Figure 2 shows approximate Fe concentration plotted against the intensity of carbon K radiation in four Dimmitt inclusions. The data define a mixing curve between C and Fe_3O_4. Strong absorption of carbon X rays by Fe causes curvature of the mixing line. Table 2 shows mean compositions of graphite–magnetite inclusions in Dimmitt

Table 1 X-ray powder powder diffraction patterns of a graphite–magnetite inclusion in the Dimmitt chondrite*, and of graphite, magnetite and aragonite standards

Dimmitt inclusion†		Graphite§ C (ref. 38)		Magnetite Fe_3O_4 (ref. 39)		Aragonite $CaCO_3$ (ref. 39)	
d(Å)	I/I_0	*d*(Å)	I/I_0	*d*(Å)	I/I_0	*d*(Å)	I/I_0
4.21‡	vw						
3.39	vs	3.37	10			3.396	10
3.28	w					3.273	5
2.956	vw			2.967	3		
2.700	w					2.700	5
2.520	m			2.532	10		
2.481	vw					2.481	3
2.383	vw					2.372	4
2.337	vw					2.341	3
2.106	w	2.136	4	2.099	2	2.106	2
2.026	vw	2.044	5				
1.977	w					1.977	7
1.882	vw					1.882	3
						1.877	3
1.815	vw					1.814	2
1.740	vw					1.742	3
1.689	vw	1.691	5				
1.605	vw			1.616	3		
1.477	vw			1.485	4		

* Using FeKα radiation.

† After treatment with 6M HCl the sample gave only graphite powder lines 3.39 (vs), 2.203 (w), 1.690 (vw), 1.288 (vw) and 1.155 (vw). vs, Very strong; w, weak; vw, very weak; m, medium.

‡ Probably the strongest line of goethite (α-FeOOH).

§ Spacings and intensities of graphite diffraction lines vary considerably.

---

* Permanent address: Institute of Physical and Chemical Research, Wako, Saitama, Japan.

0028-0836/81/250544—03$01.00

and two other meteorites obtained with a 10-μm wide electron beam. Even with such a broad beam, elemental concentrations vary considerably from place to place within a single inclusion: Fe usually varies by 10–50% relative from the mean, minor elements such as Si, Ni, S, Cr and Mg vary by ~50–100% relative, although their total is usually below 5 wt%. Carbon concentrations, which vary overall between 40 and 98 wt%, were estimated by difference assuming that S is in the form of FeS, the remaining Fe is in Fe_3O_4 and other elements are present as oxides. Analyses of graphite–magnetite in seven other chondrites all plot on or near the mixing curve shown in Fig. 2, and show comparable minor element concentrations.

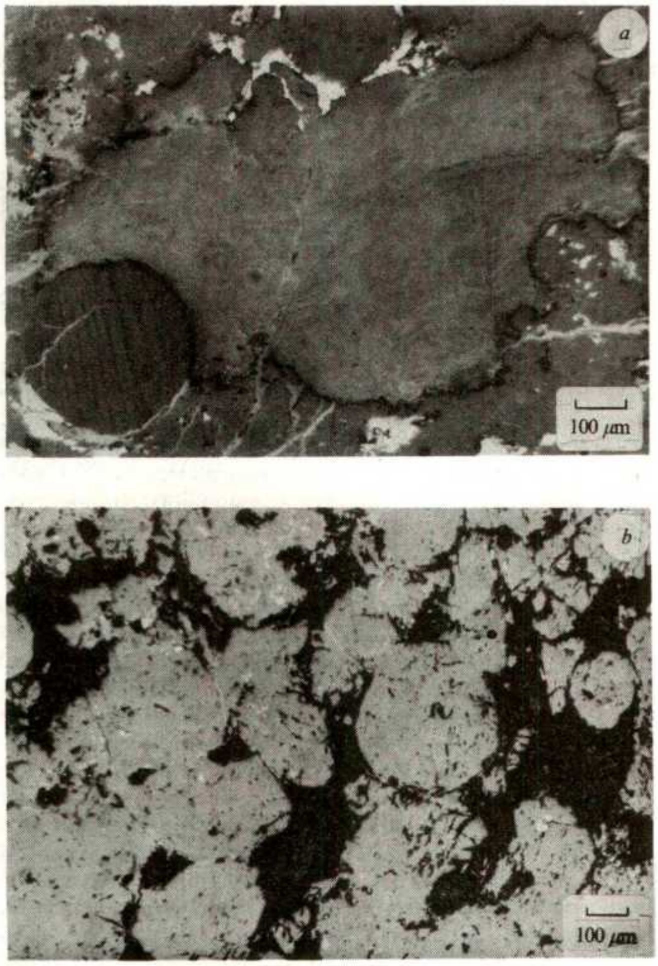


Fig. 1 Reflected-light photomicrographs of graphite–magnetite aggregates in ordinary chondrites. *a*, An irregular isolated graphite–magnetite inclusion in the Dimmitt regolith breccia. (Veins were produced by terrestrial weathering.) *b*, Part of the clast PV1 in the Plainview regolith breccia. This clast is a new kind of chondrite that contains chondrules in a graphite–magnetite matrix.

Isolated graphite–magnetite inclusions like those in Dimmitt were also found in six other H chondrites—Abbott, Elm Creek, Leighton, Plainview, Tysnes Island and Weston—and in one L chondrite—Grassland. All these meteorites contain abundant solar-wind gases[5,6], indicating that they are regolith breccias[7,8] which formed by collisional mixing of meteorite material. We estimate that these chondrites contain between 10^{-2} and 10^{-6} g per g of graphite–magnetite inclusions, and predict that other gas-rich ordinary chondrites, and perhaps gas-rich achondrites as well, contain similar concentrations of inclusions.

Table 2 Mean compositions (wt %)* of graphite–magnetite in Dimmitt, Plainview and Allan Hills A77011 chondrites determined by broad-beam electron-microprobe analysis (5–10 μm beam diameter)

Meteorite	Dimmitt			Plainview	A77011‡
No. of analyses	Incl. 1†	Incl. 5	Clast DT1	Clast PV1	
	20	4	24	28	12
Na	0.09	0.13	0.18	0.08	<0.05
Mg	<0.03	0.11	0.4	0.6	0.4
Al	<0.03	<0.03	0.15	0.08	0.06
Si	0.35	0.7	1.4	1.4	0.6
P	<0.03	<0.03	0.1	0.04	<0.03
S	0.11	0.4	0.5	0.25	0.25
Ca	0.09	0.06	0.15	0.4	0.09
Cr	<0.05	<0.05	0.2	0.5	<0.05
Fe	27	9	31	11	9
Ni	1.0	0.4	4	3	1.4
C by diff.	(60)	(84)	(47)	(75)	(83)

* Other elements measured: Cl ≤ 0.1%, K < 0.1%, Ti < 0.05%, Mn ≤ 0.1%, Co < 0.1% (0.6% in DT1), Cu < 0.2%, and Zn < 0.3%.

† Inclusion is shown in Fig. 1*a*, and plotted in Fig. 2 as inclusion 1.

‡ Specimen Allan Hills A77043, which like A77214 and several other Allan Hills specimens is paired with Allan Hills A77011.

A new kind of unequilibrated ordinary chondrite containing even higher concentrations of graphite–magnetite was discovered in the form of four clasts in Plainview (PV1, which was described by Fodor and Keil[9]), Dimmitt (DT1 and DT2) and Weston (WN1). These four clasts, which are only 1–5 mm in diameter, contain 15–35 vol% of graphite–magnetite matrix, which fills much of the space between the chondrules (Fig. 1*b*). Like normal unequilibrated ordinary chondrites (UOCs), the clasts contain heterogeneous olivines and pyroxenes, olivines with high Ca concentrations, and well-defined chondrules containing glass[1]. But they differ in having a graphite–magnetite matrix instead of the normal Huss matrix, and consequently have much higher C concentrations (7–17 wt%) than normal UOCs (0.2–1.0 wt%) and other meteorites[10,11]. Metallic Fe,Ni abundances in the clasts seem to be lower than those of most UOCs. However, shock veining and terrestrial corrosion prevent accurate estimates of metal abundances.

We have discovered two UOCs that contain abundant graphite–magnetite in addition to the normal Huss matrix: Sharps (H3) and Allan Hills A77011 (L3). Sharps contains more graphite–magnetite than Huss matrix, whereas in A77011, the

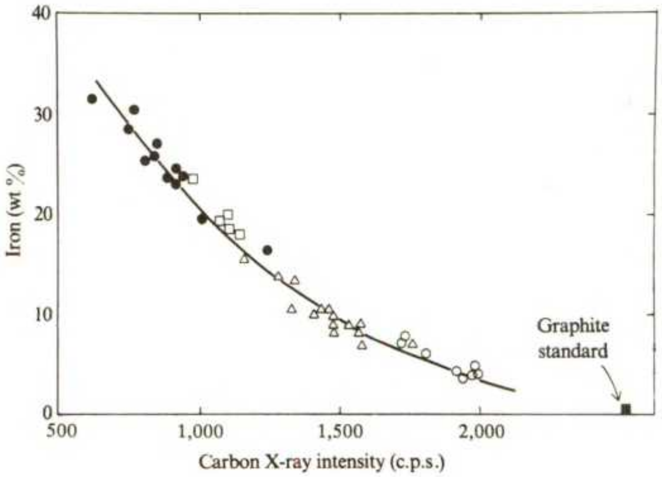


Fig. 2 Approximate Fe concentration in four Dimmitt inclusions (●, 1; □, 2; △, 3; ○, 4) plotted against the intensity of carbon X rays (K line, background corrected) measured simultaneously with the electron microprobe. The data define a mixing curve between pure carbon and an Fe-rich mineral, consistent with X-ray diffraction data, which indicate that the inclusions are largely composed of graphite and magnetite.

reverse is true. These two meteorites contain very heterogeneous silicates and have high bulk carbon concentrations[12,13], characteristics of the most unequilibrated ordinary chondrites[1,3,14]. In both meteorites, tiny metallic Fe,Ni grains commonly occur within graphite–magnetite inclusions. Their morphology and distribution suggest that they formed by reduction of magnetite, possibly before the meteorites accreted. (Fredriksson *et al.*[12] also observed graphite in Sharps, but not magnetite, and suggested that metal formed by the reducing action of carbon.) Four other UOCs have been found which contain very minor graphite–magnetite (<1 vol%): Allan Hills A77278 (LL3), Allan Hills A77299 (H3), Dhajala (H3) and Bremervörde (H3). In these meteorites also, the graphite–magnetite is associated with Fe,Ni metal and distributed around chondrules. A full description of these clasts and chondrites will be given elsewhere.

Discussion

Although composed of very abundant elements, graphite and magnetite have been considered relatively rare minerals in ordinary chondrites[15]. Their aggregates may have escaped previous detection because in transmitted light they appear opaque like metal or unrecrystallized Huss matrix, whereas in reflected light they look grey like silicates. Here we discuss the origin and implications of graphite–magnetite and its occurrence in unequilibrated clasts and meteorites and in chondritic regolith breccias.

The presence of neon-E, which is almost pure ^{22}Ne (ref. 16), in a graphite-rich acid-resistant residue from Dimmitt[17] suggests that the graphite–magnetite we observed might be the carrier of Ne-E and, hence, presolar in origin[18,19]. (See also ref. 20 for views on the origin of Ne-E.) However, noble-gas analyses of samples from a large graphite–magnetite inclusion in Dimmitt and the clast PV1 failed to reveal Ne-E (ref. 21). Thus, we conclude that either Ne-E is heterogeneously distributed in these inclusions, or it is present in another phase in the residue. The abundance of submicrometre graphite, and perhaps magnetite also, in interstellar grains[22,23] and the occurrence of Ne-E in carbon-rich phases in CI and CM carbonaceous chondrites[24-26] provide weak evidence favouring a presolar source for the inclusions. If they are presolar, their abundance in the unequilibrated clasts implies that at the formation location of these clasts in the solar nebula, a significant fraction of the accreting material was interstellar dust and not solar nebula condensates. This would support Clayton's view[27] that temperatures were not high enough to vaporize interstellar grains at the formation location of meteorites.

Like Huss matrix, the graphite–magnetite component seems to have formed by gas–solid reactions before accretion of the meteorites. According to equilibrium calculations, in a gas of solar composition, graphite only forms at low temperatures and pressures, less then 400 K and 10^{-8} atm (refs 28, 29). However, it is unlikely that equilibrium could be obtained in practice in these conditions. Alternatively, if the carbon abundance in the gas exceeds that of oxygen, graphite is stable over a wide range of temperatures and pressures and is the most abundant solid to form at equilibrium. Iron metal is stable at 10^{-4} atm below 1,400 K, but reacts to form FeS below ~600 K and Fe_3O_4 below ~400 K in systems with solar abundances[22,30]. Thus graphite and iron are major condensates in C-rich environments, but if the graphite–magnetite inclusions equilibrated with a gas it is difficult to understand their low abundances of FeS and Si-rich minerals. Huss matrix, which has been considered to be a low-temperature condensate[3,30], also contains C and Fe_3O_4 (refs 31, 32) but in concentrations far lower than those of the inclusions we have described.

The graphite–magnetite component resembles the Huss matrix in its submicrometre crystal size, and circumstantial evidence suggests that both may be enriched in volatile elements[3,21]. However, we have not observed any mixing of these components and conclude that they formed in very different conditions. We suggest that UOCs formed from four major components: chondrules, metallic Fe,Ni, Huss matrix, and graphite–magnetite. For unknown reasons, some UOC's failed to accrete Huss matrix and are entirely composed of chondrules and graphite–magnetite. We conclude that this material, represented by the four clasts, constitutes a new kind of unequilibrated chondrite.

Finally, we suggest that unequilibrated chondrites rich in graphite–magnetite were the source of the abundant graphite–magnetite inclusions in ordinary-chondritic regolith breccias. These chondrites may have been derived from subkilometre bodies that were too small to be metamorphosed by ^{26}Al heating. (Chondritic regolith breccias and unequilibrated chondrites experienced only mild metamorphism after compaction. Stronger heating might have reduced magnetite to metal.) These unequilibrated chondrites may also have been a source for the carbon (and volatile?) enrichments in the dark hosts of ordinary-chondritic regolith breccias[33,34]. Neither carbonaceous chondrites nor normal UOCs seem capable of providing the observed volatile enrichments[34,35]. A flux of bodies rich in graphite–magnetite in the Solar System might also account for some of the discrepancies between the reflectivities of asteroids and common chondrites[36]. Many asteroids have flat reflectance spectra and low albedoes, characteristics that can be reproduced by the presence of a small admixture of opaques like graphite and magnetite[37].

We thank R. S. Clarke Jr, M. E. Lipschutz and the Antarctic Meteorite Working Group for the loan of meteorite samples, G. H. Conrad and J. S. Allen for technical assistance, R. Hutchison for comments, and B. Hudson and C. M. Hohenberg for noble-gas analyses. This work was supported by NASA grant NGL 32-004-064.

Received 27 January; accepted 1 April 1981.

1. Dodd, R. T., Van Schmus, W. R. & Kaufman, D. M. *Geochim. cosmochim. Acta* **31,** 921–951 (1967).
2. Wood, J. A. *Icarus* **2,** 152–180 (1963).
3. Huss, G. R., Keil, K. & Taylor, G. J. *Geochim. cosmochim. Acta* **45,** 33–51 (1981).
4. Reed, S. J. B. *Electron Microprobe Analysis,* 318–325 (Cambridge University Press, 1975).
5. Schultz, L. & Kruse, H. *Nucl. Track Detect.* **2,** 65–103 (1975).
6. Wasson, J. T. *Meteorites—Classification and Properties,* 97–104 (Springer, Berlin, 1974).
7. Pellas, P. in *From Plasma to Planet* (ed. Elvius, A.) 65–92 (Wiley, New York, 1972).
8. Rajan, R. S. *Geochim. cosmochim. Acta* **38,** 777–788 (1974).
9. Fodor, R. V. & Keil, K. *Geochim. cosmochim. Acta* **40,** 177–189 (1976).
10. Moore, C. B. & Lewis, C. F. *J. geophys. Res.* **72,** 6289–6292 (1967).
11. Otting, W. & Zähringer, J. *Geochim. cosmochim. Acta* **31,** 1949–1960 (1967).
12. Fredriksson, K., Jarosewich, E. & Nelen, J. in *Meteorite Research* (ed. Millman, P. M.) 155–165 (Springer, Berlin, 1969).
13. Marvin, U. B. & Mason, B. *Smiths. Contr. Earth Sci.* **23,** 1–50 (1980).
14. Sears, D. W., Grossman, J. N., Melcher, C. L., Ross, L. M. & Mills A. A. *Nature* **287,** 791–795 (1980).
15. Ramdohr, P. *The Opaque Minerals in Stony Meteorites* (Elsevier, Amsterdam, 1973).
16. Jungck, M. H. A. & Eberhardt, P. *Meteoritics* **14,** 439–441 (1979).
17. Niederer, F. & Eberhardt, P. *Meteoritics* **12,** 327–331 (1977).
18. Black D. C. *Geochim. cosmochim. Acta* **36,** 377–394 (1972).
19. Clayton, D. D. *Nature* **257,** 36–37 (1979); *Space Sci. Rev.* **24,** 147–226 (1979).
20. Sabu, D. D. & Manuel, O. K. *Proc. 11th Lunar Planet. Sci. Conf.* 879–899 (1980).
21. Scott, E. R. D. *et al. Lunar planet Sci.* **12,** 955–957 (1981).
22. Salpeter, E. E. *A. Rev. Astr. Astrophys.* **15,** 267–293 (1977).
23. Knacke, R. F. in *Protostars and Planets* (ed. Gehrels, T.) 112–133 (University of Arizona, Tucson, 1978).
24. Lewis, R. S., Alaerts, L., Matsuda, J. & Anders, E. *Astrophys. J. Lett.* **234,** L165–L168 (1979).
25. Eberhardt, P., Jungck, M. H. A., Meier, F. O. & Niederer F. *Astrophys. J. Lett.* **234,** L169–L171 (1979).
26. Whittaker, A. G., Watts, E. J., Lewis, R. S. & Anders, E. *Science* **209,** 1512–1514 (1980).
27. Clayton, D. D. *Moon Planets* **19,** 109–137 (1978).
28. Sears, D. W. *Earth planet. Sci. Lett.* **41,** 128–138 (1978).
29. Lewis, J. S., Barshay, S. S. & Noyes, B. *Icarus* **37,** 190–206 (1979).
30. Larimer, J. W. *Astrophys. Space Sci.* **65,** 351–369 (1979).
31. Kurat, G. *Earth planet. Sci. Lett.* **7,** 317–324 (1970).
32. Wood J. A. *Geochim. cosmochim. Acta* **26,** 739–749 (1962).
33. Müller, O. & Zähringer, J. *Earth planet. Sci. Lett.* **1,** 25–29 (1966).
34. Rieder, R. & Wänke, H. in *Meteorite Research* (ed. Millman, P. M.) 75–86 (Reidel, Dordrecht, 1969).
35. Bart, G. & Lipschutz, M. E. *Geochim. cosmochim. Acta* **43,** 1499–1504 (1979).
36. Gaffey, M. J. & McCord, T. B. in *Asteroids* (ed. Gehrels, T.) 688–723 (1979).
37. Johnson, T. V. & Fanale, F. P. *J. geophys. Res.* **85,** 8507–8518 (1973).
38. Vdovykin, G. P. *Space Sci. Rev.* **14,** 758–831 (1973).
39. JCPDS International Center for Diffraction Data, Card NO's 19–629 and 5–453 (1980).

Primary structure, gene organization and polypeptide expression of poliovirus RNA

Naomi Kitamura*, Bert L. Semler, Paul G. Rothberg, Glenn R. Larsen, Cheryl J. Adler, Andrew J. Dorner, Emilio A. Emini, Ronnie Hanecak, James J. Lee*, Sylvie van der Werf*, Carl W. Anderson† & Eckard Wimmer‡

Department of Microbiology, School of Medicine, State University of New York at Stony Brook, Stony Brook, New York 11794, USA and
† Department of Biology, Brookhaven National Laboratory, Upton, New York 11973, USA
* Present addresses: Department of Viral Oncology, Cancer Institute, 1-37-1 Kami-Ikebukuru, Toshima-ku, Tokyo 170, Japan (N.K.); Department of Biology, California Institute of Technology, Pasadena, California 91125, USA (J.J.L.); Unité de Virologie Moléculaire et Unité de Génie Génétique, Institut Pasteur, Paris, France (S.v.d.W.) ‡ To whom correspondence should be addressed

---

The primary structure of the poliovirus genome has been determined. The RNA molecule is 7,433 nucleotides long, polyadenylated at the 3′ terminus, and covalently linked to a small protein (VPg) at the 5′ terminus. An open reading frame of 2,207 consecutive triplets spans over 89% of the nucleotide sequence and codes for the viral polyprotein NCVPOO. Twelve viral polypeptides have been mapped by amino acid sequence analysis and were found to be proteolytic cleavage products of the polyprotein, cleavages occurring predominantly at Gln-Gly pairs.

POLIOVIRUS was discovered to be the causative agent of poliomyelitis by Landsteiner and Popper in 1908[1]. Although development of both killed and live vaccines[2,3] has been spectacularly successful in controlling poliomyelitis in industrialized countries since the mid 1950s, poliomyelitis remains a serious threat in much of the world. In spite of intensive research, the precise mechanisms by which poliovirus replicates and causes cytopathogenic effects, and the molecular basis for the attenuation of live virus vaccines, remain unknown.

We have now determined the primary nucleotide sequence of the entire single-stranded genome of poliovirus type 1 (Mahoney strain). Elucidation of the primary structure of the viral genome might reveal elements important to virus replication that might not have been apparent from other experimental approaches, such as the genome-linked protein VPg found attached to the 5′ ribonucleotide[4–6], the initiation and termination sites for poliovirus translation, structural signals in protein processing[7] and potential elements of secondary structure within the genome itself[8].

Genome RNA and viral mRNA are thought to be identical in nucleotide sequence and differ only in the presence or absence of VPg[9,10]. Hence, studies of genome RNA permit an analysis of gene expression by translation. Poliovirus follows a unique strategy of gene expression by synthesizing a primary polyprotein which may contain all of the virus' translatable information[11] and which is cleaved post-translationally into numerous polypeptides[11–13]. The relationships of these cleavage products to each other have been determined[14]. However, the signals directing protein processing, the proteinases involved, and the possibility of regulating viral replication through polypeptide cleavage have only been matters of speculation.

Sequence analysis of viral RNA

Poliovirus RNA was sequenced in three phases: (1) the identification of the 3′-terminal poly(A), (2) the determination of internal sequences, and (3) the characterization of a 5′-terminal, protein-linked fragment.

Following Kates'[16] discovery of polyadenylate in eukaryotic mRNA, poliovirus RNA was analysed and found to contain a poly(A) tract[17,18] at its 3′ end[18]. This structural element permitted the reverse transcription of poliovirus RNA into cDNA using oligo(dT) primers and reverse transcriptase[19,20].

Poliovirus RNA can be readily reverse-transcribed, so a logical strategy would be to clone the cDNA and then sequence the cloned DNA. Until recently this strategy of analysis was prohibited for poliovirus RNA by the NIH recombinant DNA guidelines. Instead, we sequenced poliovirus cDNA by a modification[20] of Sanger's chain termination method[21] as follows: (1) Poliovirus cDNA was synthesized and chains of 7,000–7,400 deoxyribonucleotides were selected by zonal centrifugation. (2) Poliovirion RNA was digested exhaustively with RNase T_1 or with RNase A, and the large oligonucleotide products were separated by two-dimensional gel electrophoresis. Large oligonucleotides ($n > 10$) were eluted from the gel, dephosphorylated at their 3′ ends and labelled at their 5′ ends with phosphorus-32. (3) In four reaction mixtures the cDNA and a specific, 5′-^{32}P-labelled primer oligonucleotide were annealed and then incubated with *Escherichia coli*

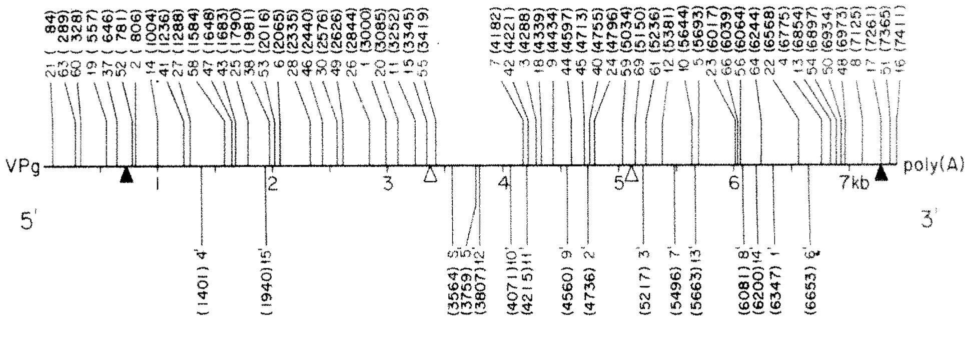


Fig. 1 Physical map of RNase T_1- (upper row) and RNase A-resistant (lower row) oligonucleotides ($n > 10$) of poliovirus RNA. Numbers refer to oligonucleotides as identified by two-dimensional polyacrylamide gel electrophoresis[20,24]. The location, within the RNA sequence, of the 5′-terminal nucleotide of an oligonucleotide is given in parentheses. With these locations the sequences of the oligonucleotides can be obtained from Fig. 2. With the exception of numbers 66, 56, 64, 54, 17, 51, 16, and 11′, all oligonucleotides were used in the chain termination method. Note that oligonucleotides 19, 47 and 4′ were poor primers, and 55, 15, 5′ and 13′ did not prime DNA synthesis at all. 'S' refers to a synthetic decadeoxyribonucleotide (Collaborative Research) used as a primer to complete the sequence. ▲, Suggested sites of initiation and termination of translation; △, position of the primary cleavage sites of the polyprotein (see Fig. 4).

0028–0836/81/250547—07$01.00

```
               10                                      40                                      70                                      100
VPg-UU AAA ACA GCU CUG GGG UUG UAC CCA CCC CAG AGG CCC ACG UGG CGG CUA GUA CUC CGG UAU UGC GGU ACC CUU GUA CGC CUG UUU UAU ACU CCC UUC CCG
                                    130                                     160                                     190                                     220
END                                                                                                         END                                     END
UAA CUU AGA CGC ACA AAA CCA AGU UCA AUA GAA GGG GGU ACA AAC CAG UAC CAC CAC GAA CAA GCA CUU CUG UUU CCC CGG UGA UGU CGU AUA GAC UGC UUG CGU GGU UGA AAG CGA CGG
                                    250                                     280                                     310                                     340
                    MET                                                 MET                                     END             MET
AUC CGU UAU CCG CUU AUG UAC UUC GAG AAG CCC AGU ACC ACC UCG GAA UCU UCG AUG CGU UGG UUA GCA CUC AAC CCC AGA GUG UAG CUU AGG CUG AUG AGU CUG GAC AUC CCU CAC CGG
                                    370                                     400                                     430                                     460
END                                                 END     MET         END     END
UGA CGG UGU UCC AGG CUG CGU UGG CGG CCU ACC UAU GGC UAA CGC AUG GGA CGC UAG UUG UGA ACA AGG UGU GAA GAG CCU AUU GAG CUA CAU AAG AAU CCU CCG GCC CCU GAA UGC GGC
                                    490                                     520                                     550                                     580
END                                             END
UAA UCC CAA CCU CGG AGC AGG UGG UUC ACA AAC CAG UGA UUG GCC UGU CGU AAC GCG CAA GUC CGU GGC GGA ACC GAC UAC UUU GGG UGU CCG UGU UUC CUU UUA UUU UAU UGU GGC UGC
                                    610                                     640                                     670                                     700
        END                                                             END     END                                                 END                 END VAL
UUA UGG UGA CAA UCA CAG AUU GUU AUC AUA AAG CGA AUU GGA UUG GCC AUC CGG UGA AAG UGA GAC UCA UUA UCU AUC UGU UUG CUG GAU CCG CUC CAU UGA GUG UGU UUA CUC UAA GUA
                                    730             ~~~→VP4         760                                     790                                     820
GLN PHE GLN GLN LEU PHE GLN SER ASP ASN CYS ILE ILE MET GLY ALA GLN VAL SER SER GLN LYS VAL GLY ALA HIS GLU ASN SER ASN ARG ALA TYR GLY GLY SER THR ILE ASN TYR
CAA UUU CAA CAG UUA UUU CAA UCA GAC AAU UGU AUC AUA AUG GGU GCU CAG GUU UCA UCA CAG AAA GUG GGC GCA CAU GAA AAC UCA AAU AGA GCG UAU GGU GGU UCU ACC AUU AAU UAC
                                    850                                     880                                     910                                     940
THR THR ILE ASN TYR TYR ARG ASP SER ALA SER ASN ALA ALA SER LYS GLN ASP PHE SER GLN ASP PRO SER LYS PHE THR GLU PRO ILE LYS ASP VAL LEU ILE LYS THR ALA PRO MET
ACC ACC AUU AAU UAU UAU AGA GAU UCA GCU AGU AAC GCG GCU UCG AAA CAG GAC UUC UCU CAA GAC CCU UCC AAG UUC ACC GAG CCC AUC AAG GAU GUC CUG AUA AAA ACA GCC CCA AUG
        →VP2                        970                                     1000                                    1030                                    1060
LEU ASN SER PRO ASN ILE GLU ALA CYS GLY TYR SER ASP ARG VAL LEU GLN LEU THR LEU GLY ASN SER THR ILE THR THR GLN GLU ALA ALA ASN SER VAL VAL ALA TYR GLY ARG TRP
CUA AAC UCG CCA AAC AUA GAG GCU UGC GGG UAU AGC GAU AGA GUA CUG CAA UUA ACA CUG GGA AAC UCC ACU AUA ACC ACA CAG GAG GCG GCU AAU UCA GUA GUC GCU UAU GGG CGU UGG
                                    1090                                    1120                                    1150                                    1180
PRO GLU TYR LEU ARG ASP SER GLU ALA ASN PRO VAL ASP GLN PRO THR GLU PRO ASP VAL ALA ALA CYS ARG PHE TYR THR LEU ASP THR VAL SER TRP THR LYS GLU SER ARG GLY TRP
CCU GAA UAU CUG AGG GAC AGC GAA GCC AAU CCA GUG GAC CAG CCG ACA GAA CCA GAC GUC GCU GCA UGC AGG UUU UAU ACG CUA GAC ACC GUG UCU UGG ACG AAA GAG UCG CGA GGG UGG
                                    1210                                    1240                                    1270                                    1300
TRP TRP LYS LEU PRO ASP ALA LEU ARG ASP MET GLY LEU PHE GLY GLN ASN MET TYR TYR HIS TYR LEU GLY ARG SER GLY TYR THR VAL HIS VAL GLN CYS ASN ALA SER LYS PHE HIS
UGG UGG AAG UUG CCU GAU GCA CUG AGG GAC AUG GGA CUC UUU GGG CAA AAU AUG UAC UAC CAC UAC CUA GGU AGG UCC GGG UAC ACC GUG CAU GUA CAG UGU AAC GCC UCC AAA UUC CAC
                                    1330                                    1360                                    1390                                    1420
GLN GLY ALA LEU GLY VAL PHE ALA VAL PRO GLU MET CYS LEU ALA GLY ASP SER ASN THR THR THR MET HIS THR SER TYR GLN ASN ALA ASN PRO GLY GLU LYS GLY GLY THR PHE THR
CAG GGG GCA CUA GGG GUA UUC GCC GUA CCA GAG AUG UGU CUG GCC GGG GAU AGC AAC ACC ACU ACC AUG CAC ACC AGC UAU CAA AAU GCC AAU CCU GGC GAG AAA GGA GGC ACU UUC ACG
                                    1450                                    1480                                    1510                                    1540
GLY THR PHE THR PRO ASP ASN ASN GLN THR SER PRO ALA ARG SER SER ALA ARG TRP ILE THR SER LEU GLU MET ALA ARG CYS TRP GLY MET PRO LEU CYS SER ALA GLN ILE ILE ASN
GGU ACG UUC ACU CCU GAC AAC AAC CAG ACA UCA CCU GCC CGC AGU UCU GCC CGG UGG AUU ACC UCC UUG GAA AUG GCA CGU UGU UGG GGA AUG CCU UUG UGU UCC GCA CAG AUA AUA AAC
                                    1570                                    1600                                    1630                                    1660
LEU ARG THR ASN ASN CYS ALA THR LEU VAL LEU PRO TYR VAL ASN SER LEU SER LEU ASP SER MET VAL LYS HIS ASN ASN TRP GLY ILE ALA ILE LEU PRO LEU ALA PRO LEU ASN PHE
UUA CGG ACC AAC AAC UGU GCU ACA CUG GUA CUA CCU UAC GUG AAC UCC CUC UCG CUA GAU AGU AUG GUA AAG CAC AAU AAU UGG GGA AUU GCA AUA UUA CCA UUG GCC CCA UUA AAU UUU
                                    1690                                    1720                                    1750            →VP3                1780
VAL SER GLU SER SER PRO GLU ILE PRO ILE THR LEU THR ILE ALA PRO MET CYS CYS GLU PHE ASN GLY LEU ARG ASN ILE THR LEU PRO ARG LEU GLN GLY LEU PRO VAL MET ASN THR
GUU AGU GAG UCC UCC CCA GAG AUU CCA AUC ACC UUG ACC AUA GCC CCU AUG UGC UGU GAG UUC AAU GGA UUA AGA AAC AUC ACC CUG CCA CGC UUA CAG GGC CUG CCG GUC AUG AAC ACC
                                    1810                                    1840                                    1870                                    1900
PRO GLY SER ASN GLN TYR LEU THR ALA ASP ASN PHE GLN SER PRO CYS ALA LEU PRO GLU PHE ASP VAL THR PRO PRO ILE ASP ILE PRO GLY GLU VAL LYS ASN MET MET GLU LEU ALA
CCU GGU AGC AAU CAA UAU CUU ACU GCA GAC AAC UUC CAG UCA CCG UGU GCG CUG CCU GAA UUU GAU GUG ACC CCA CCU AUU GAC AUA CCC GGU GAA GUA AAG AAC AUG AUG GAA UUG GCA
                                    1930                                    1960                                    1990                                    2020
GLU ILE ASP THR MET ILE PRO PHE ASP LEU SER ALA THR LYS LYS ASN THR MET GLU MET TYR ARG VAL ARG LEU SER ASP LYS PRO HIS THR ALA ALA SER ILE LEU CYS LEU SER LEU
GAA AUC GAC ACC AUG AUU CCC UUU GAC UUA AGU GCC ACA AAA AAG AAC ACC AUG GAA AUG UAU AGG GUU CGG UUA AGU GAC AAA CCA CAU ACA GCC GCC UCC AUA CUC UGC CUG UCA CUC
                                    2050                                    2080                                    2110                                    2140
SER PRO ALA SER ASP PRO ARG LEU SER HIS THR MET LEU GLY GLU ILE LEU ASN TYR TYR THR HIS TRP ALA GLY SER LEU LYS PHE THR PHE LEU PHE CYS GLY SER MET MET ALA THR
UCU CCA GCC UCA GAU CCU AGG UUG UCA CAU ACU AUG CUU GGA GAA AUC CUA AAU UAC UAC ACA CAC UGG GCA GGA UCC CUG AAG UUC ACG UUU CUG UUC UGU GGA UCC AUG AUG GCA ACU
                                    2170                                    2200                                    2230                                    2260
GLY LYS LEU LEU VAL SER TYR ALA PRO PRO GLY ALA ASP PRO PRO LYS LYS ARG LYS GLU ALA MET LEU GLY THR HIS VAL ILE TRP ASP ILE GLY LEU GLN SER SER CYS THR MET VAL
GGC AAA CUG UUG GUG UCA UAC GCG CCU CCU GGA GCC GAC CCA CCA AAG AAG CGU AAG GAG GCG AUG UUG GGA ACA CAU GUG AUC UGG GAC AUA GGA CUG CAG UCC UCA UGU ACU AUG GUA
                                    2290                                    2320                                    2350                                    2380
VAL PRO TRP ILE SER ASN SER THR TYR ARG GLN THR ILE ASP ASP SER PHE THR GLU GLY GLY TYR ILE SER VAL PHE TYR GLN THR ARG ILE VAL VAL PRO LEU SER THR PRO ARG GLU
GUG CCA UGG AUU AGC AAC AGC ACG UAU CGG CAA ACC AUA GAU GAU AGU UUC ACC GAA GGC GGA UAC AUC AGC GUC UUC UAC CAA ACU AGA AUA GUC GUC CCU CUU UCG ACA CCC AGA GAG
                                    2410                                    2440                                    2470    →VP1                        2500
MET ASP ILE LEU GLY PHE VAL SER ALA CYS ASN ASP PHE SER VAL ARG LEU LEU ARG ASP THR THR HIS ILE GLU GLN LYS ALA LEU ALA GLN GLY LEU GLY GLN MET LEU GLU SER MET
AUG GAC AUC CUU GGU UUU GUG UCA GCG UGU AAU GAC UUC AGC GUG CGC UUG UUG CGA GAU ACC ACA CAU AUA GAG CAA AAA GCG CUA GCA CAG GGG UUA GGU CAG AUG CUU GAA AGC AUG
                                    2530                                    2560                                    2590                                    2620
ILE ASP ASN THR VAL ARG GLU THR VAL GLY ALA ALA THR SER ARG ASP ALA LEU PRO ASN THR GLU ALA SER GLY PRO THR HIS SER LYS GLU ILE PRO ALA LEU THR ALA VAL GLU THR
AUU GAC AAC ACA GUC CGU GAA ACG GUG GGG GCG GCA ACA UCU AGA GAC GCU CUC CCA AAC ACU GAA GCC AGU GGA CCA ACA CAC UCC AAG GAA AUU CCG GCA CUC ACC GCA GUG GAA ACU
                                    2650                                    2680                                    2710                                    2740
GLY ALA THR ASN PRO LEU VAL PRO SER ASP THR VAL GLN THR ARG HIS VAL VAL GLN HIS ARG SER ARG SER GLU SER SER ILE GLU SER PHE PHE ALA ARG GLY ALA CYS VAL THR ILE
GGG GCC ACA AAU CCA CUA GUC CCU UCU GAU ACA GUG CAA ACC AGA CAU GUU GUA CAA CAU AGG UCA AGG UCA GAG UCU AGC AUA GAG UCU UUC UUC GCG CGG GGU GCA UGC GUG ACC AUU
                                    2770                                    2800                                    2830                                    2860
MET THR VAL ASP ASN PRO ALA SER THR THR ASN LYS ASP LYS LEU PHE ALA VAL TRP LYS ILE THR TYR LYS ASP THR VAL GLN LEU ARG ARG LYS LEU GLU PHE PHE THR TYR SER ARG
AUG ACC GUG GAU AAC CCA GCU UCC ACC ACG AAU AAG GAU AAG CUA UUU GCA GUG UGG AAG AUC ACU UAU AAA GAU ACU GUC CAG UUA CGG AGG AAA UUG GAG UUC UUC ACC UAU UCU AGA
                                    2890                                    2920                                    2950                                    2980
PHE ASP MET GLU LEU THR PHE VAL VAL THR ALA ASN PHE THR GLU THR ASN ASN GLY HIS ALA LEU ASN GLN VAL TYR GLN ILE MET TYR VAL PRO PRO GLY ALA PRO VAL PRO GLU LYS
UUU GAU AUG GAA CUU ACC UUU GUG GUU ACU GCA AAU UUC ACU GAG ACU AAC AAU GGG CAU GCC UUA AAU CAA GUG UAC CAA AUU AUG UAC GUA CCA CCA GGC GCU CCA GUG CCC GAA AAA
                                    3010                                    3040                                    3070                                    3100
TRP ASP ASP TYR THR TRP GLN THR SER SER ASN PRO SER ILE PHE TYR THR TYR GLY THR ALA PRO ALA ARG ILE SER VAL PRO TYR VAL GLY ILE SER ASN ALA TYR SER HIS PHE TYR
UGG GAC GAC UAC ACA UGG CAA ACC UCA UCA AAU CCA UCA AUC UUU UAC ACC UAC GGG ACA GCU CCA GCC CGG AUC UCG GUA CCG UAU GUU GGU AUU UCG AAC GCC UAU UCA CAC UUU UAC
                                    3130                                    3160                                    3190                                    3220
ASP GLY PHE SER LYS VAL PRO LEU LYS ASP GLN SER ALA ALA LEU GLY ASP SER LEU TYR GLY ALA ALA SER LEU ASN ASP PHE GLY ILE LEU ALA VAL ARG VAL VAL ASN ASP HIS ASN
GAC GGU UUU UCC AAA GUA CCA CUG AAG GAC CAG UCG GCA GCA CUA GGU GAC UCC CUU UAU GGU GCA GCA UCU CUA AAU GAC UUC GGU AUU UUG GCU GUU AGA GUA GUC AAU GAU CAC AAC
                                    3250                                    3280                                    3310                                    3340
PRO THR LYS VAL THR SER LYS ILE ARG VAL TYR LEU LYS PRO LYS HIS ILE ARG VAL TRP CYS PRO ARG PRO PRO ARG GLN LEU ALA TYR TYR GLY PRO GLY VAL ASP TYR LYS ASP GLY
CCG ACC AAG GUC ACC UCC AAA AUC AGA GUG UAU CUA AAA CCC AAA CAC AUC AGA GUC UGG UGC CCG CGU CCA CCG AGG CAG CUG GCG UAC UAC GGC CCU GGA GUG GAU UAC AAG GAU GGU
                                    3370            →3b                     3400                                    3430                                    3460
THR LEU THR PRO LEU SER THR LYS ASP LEU THR THR TYR GLY PHE GLY HIS GLN ASN LYS ALA VAL TYR THR ALA GLY TYR LYS ILE CYS ASN TYR HIS LEU ALA THR GLN ASP ASP LEU
ACG CUU ACA CCC CUC UCC ACC AAG GAU CUG ACC ACA UAU GGA UUC GGA CAC CAA AAC AAA GCG GUG UAC ACU GCA GGU UAC AAA AUU UGC AAC UAC CAC UUG GCC ACU CAG GAU GAU UUG
                                    3490                                    3520                                    3550                                    3580
GLN ASN ALA VAL ASN VAL MET TRP SER ARG ASP LEU LEU VAL THR GLU SER ARG ALA GLN GLY THR ASP SER ILE ALA ARG CYS ASN CYS ASN ALA GLY VAL TYR TYR CYS GLU SER ARG
CAA AAC GCA GUG AAC GUC AUG UGG AGU AGA GAC CUC UUA GUC ACA GAA UCA AGA GCC CAG GGC ACC GAU UCA AUC GCA AGG UGC AAU UGC AAC GCA GGG GUG UAC UAC UGC GAG UCU AGA
                                    3610                                    3640                                    3670                                    3700
ARG LYS TYR TYR PRO VAL SER PHE VAL GLY PRO THR PHE GLN TYR MET GLU ALA ASN ASN TYR TYR PRO VAL ARG TYR GLN SER HIS MET LEU ILE GLY HIS GLY PHE GLU SER PRO GLY
AGG AAA UAC UAC CCA GUA UCC UUC GUU GGC CCA ACG UUC CAG UAC AUG GAG GCU AAU AAC UAU UAC CCA GUU AGG UAC CAG UCC CAU AUG CUC AUU GGC CAU GGA UUC GAA UCU CCA GGG
                                    3730                                    3760                                    3790                                    3820
ASP CYS GLY GLY ILE LEU ARG CYS HIS HIS GLY VAL ILE GLY ILE ILE THR ALA GLY GLY GLU GLY LEU VAL ALA PHE SER ASP ILE ARG ASP LEU TYR ALA TYR GLU GLU GLU ALA MET
GAU UGU GGU GGC AUA CUC AGA UGU CAC CAC GGG GUG AUA GGG AUC AUU ACU GCU GGU GGA GAA GGG UUG GUU GCA UUU UCA GAC AUU AGA GAC UUG UAU GCC UAC GAA GAA GAA GCC AUG
        →5b                         3850                                    3880                                    3910                                    3940
GLU GLN GLY ILE THR ASN TYR ILE GLU SER LEU GLY ALA ALA PHE GLY SER GLY PHE THR GLN GLN ILE SER ASP LYS ILE THR GLU LEU THR ASN MET VAL THR SER THR ILE THR GLU
GAA CAA GGC AUC ACC AAU UAC AUA GAG UCA CUU GGG GCC GCA UUU GGA AGU GGA UUU ACU CAG CAG AUU AGC GAC AAA AUA ACA GAG UUG ACC AAU AUG GUG ACC AGU ACC AUC ACU GAA
                                    3970                                    4000                                    4030                                    4060
LYS LEU LEU LYS ASN LEU ILE LYS ILE ILE SER SER LEU VAL ILE ILE THR ARG ASN TYR GLU ASP THR THR THR VAL LEU ALA THR LEU ALA LEU LEU GLY CYS ASP ALA SER PRO TRP
AAG CUA CUU AAG AAC UUG AUC AAG AUC AUA UCC UCA CUA GUU AUU AUA ACU AGG AAC UAU GAA GAC ACC ACA ACA GUG CUC GCU ACC CUG GCC CUU CUU GGG UGU GAU GCU UCA CCA UGG
                                    4090                                    →X                                      4150        *                           4180
GLN TRP LEU ARG LYS LYS ALA CYS ASP VAL LEU GLU ILE PRO TYR VAL ILE LYS GLN GLY ASP SER TRP LEU LYS LYS PHE THR GLU ALA CYS GLN ARG ALA LYS GLY LEU GLU TRP VAL
CAG UGG CUU AGA AAG AAA GCA UGC GAU GUU CUG GAG AUA CCU UAU GUC AUC AAG CAA GGU GAC AGU UGG UUG AAG AAG UUU ACU GAA GCA UGC CAA CGA GCU AAG GGC CUG GAG UGG GUG
                                    4210                                    4240                                    4270                                    4300
SER ASN LYS ILE SER LYS PHE ILE ASP TRP LEU LYS GLU LYS ILE ILE PRO GLN ALA ARG ASP LYS LEU GLU PHE VAL THR LYS LEU ARG GLN LEU GLU MET LEU GLU ASN GLN ILE SER
UCA AAC AAA AUC UCA AAA UUC AUU GAU UGG CUC AAG GAG AAA AUU AUC CCA CAA GCU AGA GAU AAG UUG GAA UUU GUA ACA AAA CUU AGA CAA CUA GAA AUG CUG GAA AAC CAA AUC UCA
                                    4330                                    4360                                    4390                                    4420
THR ILE HIS GLN SER CYS PRO SER GLN GLU HIS GLN GLU ILE LEU PHE ASN ASN VAL ARG TRP LEU SER ILE GLN SER LYS ARG PHE ALA PRO LEU TYR ALA VAL GLU ALA LYS ARG ILE
ACU AUA CAC CAA UCA UGC CCU AGU CAG GAA CAC CAG GAA AUU CUA UUC AAU AAU GUC AGA UGG UUA UCC AUC CAG UCU AAG AGG UUU GCC CCU CUU UAC GCA GUG GAA GCC AAA AGA AUA
                                    4450                                    4480                                    4510                                    4540
GLN LYS LEU GLU HIS THR ILE ASN ASN TYR ILE GLN PHE LYS SER LYS HIS ARG ILE GLU PRO VAL CYS LEU LEU VAL HIS GLY SER PRO GLY THR GLY LYS SER VAL ALA THR ASN LEU
CAG AAA CUA GAG CAU ACU AUU AAC AAC UAC AUA CAG UUC AAG AGC AAA CAC CGU AUU GAA CCA GUA UGU UUG CUA GUA CAU GGC AGC CCC GGA ACA GGU AAA UCU GUA GCA ACC AAC CUG
                                    4570                                    4600                                    4630                                    4660
ILE ALA ARG ALA ILE ALA GLU ARG GLU ASN THR SER THR TYR SER LEU PRO PRO ASP PRO SER HIS PHE ASP GLY TYR LYS GLN GLN GLY VAL VAL ILE MET ASP ASP LEU ASN GLN ASN
AUU GCU AGA GCC AUA GCU GAA AGA GAA AAC ACG UCC ACG UAC UCG CUA CCC CCG GAU CCA UCA CAC UUC GAC GGA UAC AAA CAA CAG GGA GUG GUG AUU AUG GAC GAC CUG AAU CAA AAC
                                    4690                                    4720                                    4750                                    4780
PRO ASP GLY ALA ASP MET LYS LEU PHE CYS GLN MET VAL SER THR VAL GLU PHE ILE PRO PRO MET ALA SER LEU GLU GLU LYS GLY ILE LEU PHE THR SER ASN TYR VAL LEU ALA SER
CCA GAU GGU GCG GAC AUG AAG CUG UUC UGU CAG AUG GUA UCA ACA GUG GAG UUU AUA CCA CCC AUG GCA UCC CUG GAG GAG AAA GGA AUC CUG UUU ACU UCA AAU UAC GUU CUA GCA UCC
                                    4810                                    4840                                    4870                                    4900
THR ASN SER SER ARG ILE SER PRO PRO THR VAL ALA HIS SER ASP ALA LEU ALA ARG ARG PHE ALA PHE ASP MET ASP ILE GLN VAL MET ASN GLU TYR SER ARG ASP GLY LYS LEU ASN
ACA AAC UCA AGC AGA AUU UCC CCC CCC ACU GUG GCA CAC AGU GAU GCA UUA GCC AGG CGC UUU GCG UUC GAC AUG GAC AUU CAG GUC AUG AAU GAG UAU UCU AGA GAU GGG AAA UUG AAC
                                    4930                                    4960                                    4990                                    5020
MET ALA MET ALA THR GLU MET CYS LYS ASN CYS HIS GLN PRO ALA ASN PHE LYS ARG CYS CYS PRO LEU VAL CYS GLY LYS ALA ILE GLN LEU MET ASP LYS SER SER ARG VAL ARG TYR
AUG GCC AUG GCU ACU GAA AUG UGU AAG AAC UGU CAC CAA CCA GCA AAC UUU AAG AGA UGC UGU CCU UUA GUG UGU GGU AAG GCA AUU CAA UUA AUG GAC AAA UCU UCC AGA GUU AGA UAC
```

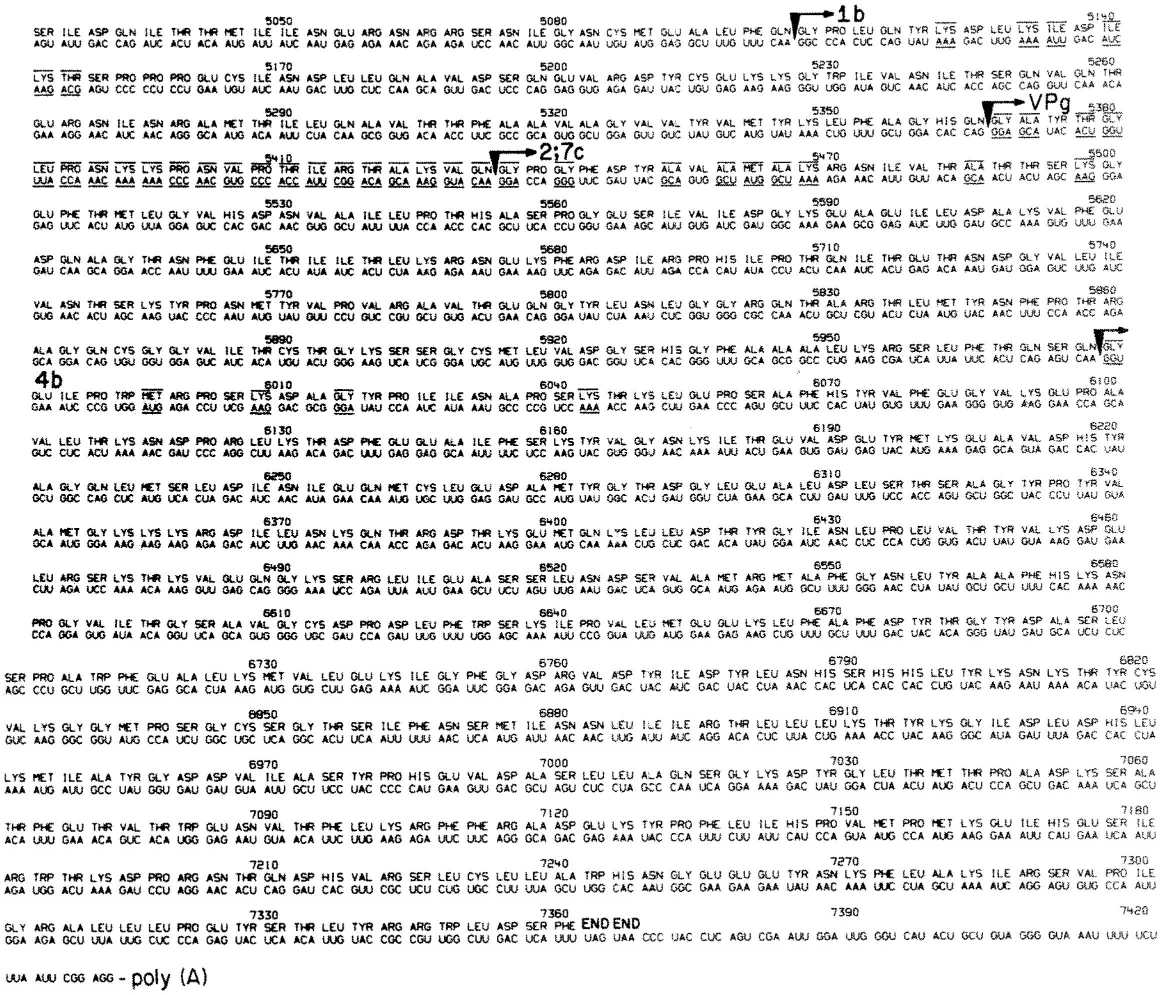


Fig. 2 The complete nucleotide sequence and encoded information of virion RNA of poliovirus type 1 (Mahoney). This strain was originally obtained from David Baltimore (Massachusetts Institute of Technology). As judged by fingerprint analyses this isolate has not undergone detectable genetic variation in our laboratory during multiple passages in HeLa cells[24]. A reading frame of 2,207 consecutive coding triplets within the nucleotide sequence was determined by computer analysis and confirmed by radiochemical microsequence analysis of 12 virus-specific polypeptides. The cleavage sites involved in processing viral polypeptides and the N-termini of viral proteins are indicated by a solid triangle and arrow; amino acids that were identified by stepwise Edman degradation of radiochemically pure polypeptides are indicated by bars. The asterisk at N4159 indicates an amino acid ambiguity. The coding sequence of capsid protein VP4 has been determined by analysis of tryptic peptides. The N-terminal amino acid of VP4 is blocked as are those of P1-1a and VPO. Preliminary evidence suggests that translation may be initiated with the methionine at N741. Termination codons in the non-coding regions are designated as 'END'. Potential initiation codons of translation in the 5'-terminal, non-coding sequence preceding N741 are indicated as MET if in phase with the major reading frame. MET codons out of phase with the major reading frame in this region are indicated with brackets. 156 nucleotides preceding the 3'-terminal poly(A) have also been determined by Porter *et al.*[66]. Analyses of the dinucleotide frequencies, codon usage and secondary structures will be published elsewhere (in preparation).

polymerase I (Klenow fragment) in the presence of unlabelled dNTPs and one of four 2',3'-dideoxynucleotide triphosphates (ddNTPs). (4) Synthesized DNA fragments (labelled through their 5'-terminal primer) were separated by slab gel electrophoresis[22] and a sequence was determined by reading the four different 'ladders' of bands. We modified the chain termination procedure of Sanger only by using radiolabelled primers and unlabelled deoxynucleotide triphosphates. This modification reduced the background of radioactivity in gels to nearly undetectable levels.

In the early stages of the investigation the newly generated sequences were ordered according to the known map positions of the largest oligonucleotides[23]. Subsequently, each newly generated sequence was searched in its 3'-terminal region for a sequence corresponding to another known, large T_1- or A-oligonucleotide[24]. This oligonucleotide was then used as a primer to extend the sequence (for details, see ref. 20).

Figure 1 shows a physical map of large ($n>10$) RNase T_1-resistant (upper row) and RNase A-resistant (lower row) oligonucleotides found in poliovirus RNA. With the exception of 'S', all of these oligonucleotides have been sequenced by rapid sequencing methods[24]. Most of them were used as primers for the chain termination method. With one exception (see below), each priming oligonucleotide produced a sequence of 250 bases, a segment long enough for overlaps with adjacent segments. Not all oligonucleotides primed DNA synthesis with equal efficiency[20], and four oligonucleotides (numbers 55, 15', 13' and 5'; ref. 24) did not prime DNA synthesis at all. Initiation at multiple sites with any single primer was not observed. 'Compressions' of bands were resolved either by slab gel electrophoresis at an elevated temperature[20] or by replacing dGTP in the reaction mixture with dITP[25] at an identical concentration, the latter being the method of choice.

The peculiar gap in T_1-oligonucleotides between nucleotides 3,419 and 4,182 (Fig. 1) may be without biological significance as this region is average in its G content. Unfortunately, the gap made it impossible to produce an overlap between N3570 and N3800. To solve this problem we obtained a custom-synthesized decadeoxyribonucleotide ['S', d(T-A-C-T-A-C-T-G-C-G), position N3564] that was selected on the basis of the known sequence generated by using T_1-oligonucleotide 15 as a primer for DNA synthesis. A reaction primed with 'S' did indeed produce a sequence that bridged this gap and yielded a continuous sequence of over 7,300 nucleotides. The 5'-terminal sequence ($n \sim 100$), however, was still unknown because the yield of cDNA strands spanning the entire genome was too low for sequence analysis.

The study of the 5′ end of poliovirion RNA proved difficult because no 5′-terminal nucleotide could be identified. Once the genome protein VPg was discovered and its covalent linkage to the 5′-terminal nucleotide established[5,6], it was possible to identify the 5′-terminal RNase T$_1$-resistant oligonucleotide (VPg-pU-U-A-A-A-A-C-A-G; refs 6,26). Moreover, Hewlett and Florkiewicz[27] succeeded in sequencing 20 nucleotides from the 5′ end by partial digestion of ^{125}I-labelled VPg-RNA. A similar approach was developed by Nomoto and Imura[28] who obtained a sequence of ~30 nucleotides (A. Nomoto, personal communication).

The remaining gap in the 5′-terminal sequence was determined using several approaches. Poliovirus is cleaved by RNase III at low ionic strength into numerous fragments[29] of which two ($n = 85$ and $n = 145$) originate from the 5′ end[30]. The smaller RNase III fragment was sequenced by rapid techniques after 3′-labelling with [5′-^{32}P]pCp, and by analysis of its T$_1$-resistant oligonucleotides[8]. The latter approach was necessary because of a stable stem-and-loop structure within the sequence that is resistant to digestion conditions of partial cleavage (see Fig. 3). The 5′-terminal sequence of the fragment was verified by analysis of poliovirus 5′-^{32}P-labelled mRNA[8]. Such analysis of the 5′ end of mRNA was possible because poliovirus mRNA does not contain VPg but has a 5′-terminal pU[31]; mRNA can therefore be dephosphorylated and end-labelled with polynucleotide kinase and [γ-^{32}P]ATP. The sequences of the large RNase III fragments and of the 5′ terminus of the mRNA connected with the sequence generated by the dideoxy method with T$_1$-oligonucleotide 21 (that maps at nucleotide 84) as the primer. This result completed the sequence analysis of poliovirus RNA.

The total nucleotide sequence of the heteropolymeric region of the genome of poliovirus type 1 (Mahoney) is shown in Fig. 2. Together with its 3′-terminal poly(A) (average chain length $n = 60$; see below), the genome RNA has a base composition of A = 30.2%, G = 22.8%, C = 23.1% and U = 23.9%, and a molecular weight of 2.411×10^6 (free acid).

We do not claim that the nucleotide sequence shown in Fig. 2 is entirely free of errors. One problem is that only one polynucleotide strand has been sequenced. Although single nucleotide assignments may require modification, the methods used make it unlikely that stretches of nucleotides have been overlooked. The following considerations lead us to accept the sequence shown in Fig. 2 as an accurate representation of the poliovirus genome: (1) The chain termination method, as we have applied it, was carried out to sequence a population of cDNA molecules rather than a single cloned DNA copy that may or may not be representative of genomic RNA. (2) All large oligonucleotides (16% of the total sequence) have been sequenced independently and by different methods[24]. (3) A restriction enzyme map[75] of cloned cDNA agrees with that derived by computer analysis of our sequence. (4) The predicted sequence of NCVPOO is consistent with the partial amino acid sequence determined for 12 viral polypeptides.

VPg-linked 5′-terminal non-coding sequence

The 5′ end of poliovirion RNA is highly unusual. Its nucleotide sequence up to the first RNase III cleavage site is shown in Fig. 3. The sequence contains a stem-and-loop structure, corresponding to nucleotides 9–36 that could be involved in processes of regulation or recognition[8]. The 5′-terminal uridine is linked, via a phosphodiester bond, to a tyrosine residue of the genome protein, VPg[5,6,32–34]. VPg is a virus-encoded[5,35,36], basic (pI > 9.4; unpublished results, and ref. 6), acid-soluble, water-insoluble oligopeptide consisting of only 22 amino acids[7,36] (Fig. 2). It maps in P3, a region of the genome RNA specifying replication proteins[7,36,37]. The O^4-(5′-uridylyl) tyrosine bond is energy rich[38], but the mechanism by which it is formed is obscure, as is the function of VPg. Since poliovirus minus strand RNA has VPg linked to the 5′-terminal poly(U)[6,39,40] and because nascent RNA strands of poliovirus replicative intermediates are VPg linked[6,40], VPg has been proposed to have a role in the initiation of RNA synthesis. In addition, or alternatively, VPg may be involved in morphogenesis (reviewed in refs 41, 42).

The data suggest that translation of the large polyprotein begins within the large open reading frame at the AUG codon located at N741. This leaves a 5′ sequence of 740 potentially untranslated nucleotides. Preceding the left-most AUG in the long, open reading frame at N741 are several other AUG codons. Specifically, these occur at N185, N237, N276, N318, N376, N387, N455, and N584 (those underlined being in phase with the major reading frame). All but three of these AUG codons are followed closely by stop codons within their reading frame. However, three of the AUG triplets (at N185, N376 and N584, underscored by dotted lines) are not in phase with the polyprotein but they could initiate synthesis of polypeptides of molecular weights 4,725, 8,425, and 7,187, respectively. The significance of these potential coding sequences is not clear. The following explanations are possible: (1) Sequencing errors occurred between N400 and N774 that disconnected the shorter reading frames from the major reading frame. This explanation is unlikely because the sequence between 400 and 793 has been corroborated with cloned cDNA of poliovirus of the same strain (S.v.d.W., E. W. and M. Girard, unpublished results). (2) Poliovirus codes for one or more small proteins independently from the large polyprotein. Such small proteins may have missed detection because of their size, their lack of internal methionines, or because of rapid degradation. Ehrenfeld and co-workers[43] have detected a small polypeptide when poliovirus RNA is translated *in vitro* in the presence of 4 mM MgCl$_2$. It remains to be seen whether this product is encoded by one of the above potential coding segments. (3) The small coding segments remain untranslated.

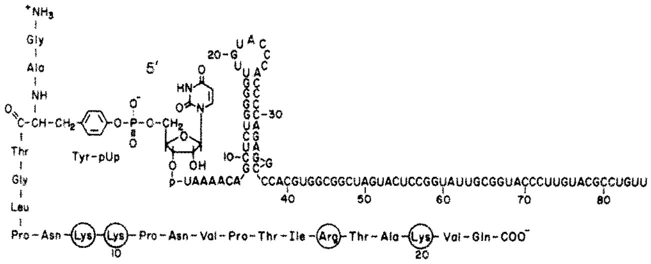

+NH3
Gly
Ala
NH
Thr
Gly
Leu
Tyr-pUp
5′
HN
O
OH
p-UAAAACA
CCACGUGGCGGCUAGUACUCCGGUAUUGCGGUACCCUUGUACGCCUGUU
10
20
30
40
50
60
70
80
Pro-Asn-Lys-Lys-Pro-Asn-Val-Pro-Thr-Ile-Arg-Thr-Ala-Lys-Val-Gln-COO−
10
20

Fig. 3 Structure of the 5′-terminus of poliovirus RNA. The oligonucleotide ($n = 85$) is an RNase III cleavage fragment of virion RNA[8]. The viral protein VPg is linked via an O^4-(5′-uridylyl)tyrosine (Tyr-pUp) bond to the RNA. The phosphodiester bond in Tyr-pUp[32,33] can be considered energy rich by analogy to O^4-(5′-adenylyl)tyrosine[38]. The stem-and-loop structure as drawn here has a $\Delta G°$ of -21 kcal mol^{-1}. The nature of the interaction between the basic genome protein (see encircled amino acids) and the RNA is unknown.

Identification of the AUG codon that functions in initiating translation of the polyprotein has not yet been possible. Assignment of the correct initiation codon is complicated by the following observations. (1) Ribosome binding experiments as carried out according to Steitz[44] have failed in this and other laboratories (E. Ehrenfeld, personal communication, and ref. 45). The major obstacle was that the ribosomes were found to bind to more than one site. Digestion of the ribosome-bound RNA with RNase yielded RNA fragments of equal length but different sequence (A. J. D. and E. W., unpublished results). (2) An equivalent to the sequence found in prokaryotic mRNA, that directs the ribosome to the initiation codon of translation[46], does not seem to exist in eukaryotic mRNA[47]. (3) The polypeptides derived from the N-terminus of the polyprotein (Pl-la, VPO and VP4) are blocked at their N-termini (unpublished results). Regardless of whether the small reading frames are translated *in vivo*, the occurrence of several AUG codons in different reading frames within this region makes it likely that some AUG codons are ignored in viral translation.

Fine structure of the protein map

In the nucleic acid sequence of poliovirus RNA (Fig. 2) an open reading frame of 2,207 consecutive triplets spans over 89% of the entire viral genome and has the coding capacity for a polypeptide of 247,000 M_r. The length of the reading frame is strong evidence that it codes for the polyprotein NCVPOO, the polypeptide proposed to represent the entire translatable information of poliovirus[11]. This conclusion is further supported by the following observations: (1) No other open reading frame longer than 78 triplets exists within the sequence. (2) All major viral polypeptides, 12 of which we have analysed by protein sequencing, map in phase within the long reading frame. These data prove that all known poliovirus proteins can originate by proteolytic cleavage from a single precursor polypeptide[11–13].

Individual poliovirus proteins were positioned with respect to the nucleotide sequence by partial amino acid sequence analysis and by alignment of the predicted with the experimental sequences. Viral polypeptides were labelled *in vivo* with one amino acid at a time, purified by polyacrylamide gel electrophoresis and their N-terminal amino acid sequences determined by radiochemical microsequence analysis[7,36]. A computer search of the poliovirus genomic coding sequence identified for each amino acid pattern a unique encoding location (see Fig. 2).

We have divided the polypeptide map into three regions (see Fig. 4): P1 corresponds to the region of capsid polypeptides and their precursors, P2 corresponds to the central portion of the genome, and P3 to a group of non-structural polypeptides including the putative proteinase P3-7c (by analogy with studies on encephalomyocarditis virus)[48,49], VPg[7,36] and the RNA-dependent RNA polymerase P3-4b[7,50,51]. Our data unambiguously show that, contrary to a previous proposal[15], P2-X maps within the centre of the viral genome.

Our strategy of locating the N-terminus of a protein with respect to the genomic sequence failed for components P1-la, VPO and VP4 because these polypeptides have blocked N-termini. Nonetheless, we could deduce limits for the sequence that encodes VP4 from a knowledge of its estimated molecular weight (corresponding to ~70 amino acids), its reported amino acid composition[52,53], and the fact that it is the only known protein that maps to the left of VP2 (reviewed in ref. 14). Within 70 amino acids to the left of VP2, only one tyrosine-containing tryptic peptide is predicted at N794-N842 (with Tyr at positions 2, 9, 14 and 15), and that was confirmed experimentally by stepwise microsequence analysis of a total trypsin digest of ^{3}H-Tyr-labelled VP4 (see Fig. 2; unpublished results). The N-terminal boundary of VP4 must be to the right of N734, as a cysteine is predicted at N732 and VP4 does not contain cysteine[52,53]. Three codons to the right of the cysteine (N732) the genomic sequence predicts a methionine codon (AUG at N741) which is the left-most methionine codon within the large open reading frame shown in Fig. 2. This AUG codon is a likely candidate for the initiation codon of the precursor polypeptides to VP4 (see below).

The next AUG codon within the open reading frame occurs at N939, only three codons to the left of the N-terminus of VP2. This methionine would be contained within the C-terminal tryptic peptide of VP4. Analysis of a tryptic digest of ^{35}S-Met-labelled VP4 by HPLC yielded a single Met-containing tryptic peptide. Stepwise degradation showed a Met in position 4 as would be expected for the predicted C-terminal tryptic peptide (unpublished results). The presence of a single Met-containing tryptic peptide in VP4 is consistent with the removal of the methionine at N741 by analogy to previously sequenced proteins[54]. We predict that the N-terminus of VP4 arises from an expected modification of the nascent polyprotein and that it has the sequence, beginning at N744, of (block)-Gly-Ala-Gln-Val-Ser-Ser-Gln-Lys-. VP4 would then contain 68 amino acids and have a predicted molecular weight of ~7,500.

Table 1 Features of poliovirus proteins

Protein	Position in nucleotide sequence	No. of amino acids	Molecular weights	Net charge
NCVPOO	741–7,361	2,207	246,530	+44
P1-1a	744–3,380	879	97,247	+12
VPO	744–1,760	339	37,352	+3
VP4	744–947	68	7,385	+2
VP2	948–1,760	271	29,985	+1
VP3	1,761–2,474	238	26,410	0
VP1	2,475–3,380	302	33,521	+9
P2-3b	3,381–5,105	575	64,953	+16
P2-5b	3,828–5,105	426	48,273	+15
P2-X	4,119–5,105	329	37,555	+15
P3-1b	5,106–7,361	752	84,234	+16
VPg	5,367–5,432	22	2,354	+4
P3-2	5,433–7,361	643	72,132	+11
P3-4b	5,979–7,361	461	52,481	+6
P3-7c	5,433–5,978	182	19,669	+5

The initiation site of translation at N741 (NCVPOO) is based on circumstantial evidence only. The number of amino acids and molecular weights are given on the assumption that C-terminal trimming of polypeptides does not occur.

Cell-free translation experiments with aphthovirus RNA have demonstrated the existence of a polypeptide (16,000 M_r) that maps in a region preceding the P1 region[55]. The results of *in vitro* translation of encephalomyocarditis (EMC) virus RNA have also been interpreted to mean that a 'leader' peptide precedes the capsid polypeptides, but direct evidence for such a leader is lacking (for a discussion, see ref. 43). On the other hand, translation of poliovirus RNA *in vitro* in the presence of labelled formyl-methionyl-tRNA has failed to produce evidence for the existence of a comparable poliovirus 'leader peptide' (E. Ehrenfeld, personal communication; our unpublished results).

The molecular weights (Table 1) calculated from the predicted amino acid sequences (Fig. 2) agree well with those deduced from SDS-polyacrylamide gel electrophoresis. The calculations are based on the assumption that no trimming of the C-termini occurs after processing. Moreover, the sequences of capsid proteins shown in Fig. 2 agree well with published data on N-terminal amino acid sequences and amino acid compositions of poliovirus type 1 (strain 1a/S3) capsid proteins[52,56], and with the amino acid composition of type 1 (Mahoney) capsid polypeptides[57]. Heterogeneity of the capsid polypeptides, as reported by Vrijsen *et al.*[56], has not been observed in our studies.

The C-terminal amino acid of polypeptide P3-2 was determined by carboxypeptidase analysis to be phenylalanine (unpublished results). This observation and the results of nucleotide sequence analyses suggest that translation of poliovirus terminates at N7361, 73 nucleotides before the beginning of the poly(A) tract. (We previously suggested that termination of translation may occur at N6874[20,42]. This result was erroneous due to three sequence ambiguities.)

The genetic map that we have derived from our protein and RNA sequence analyses is summarized in Fig. 4. It is in agreement with a map published by Rueckert and his colleagues[58]. Our data extend this information by indicating the precise location of the N-terminus of each major poliovirus proteolytic cleavage product with respect to the genomic sequence. Products of alternate cleavage pathways are not

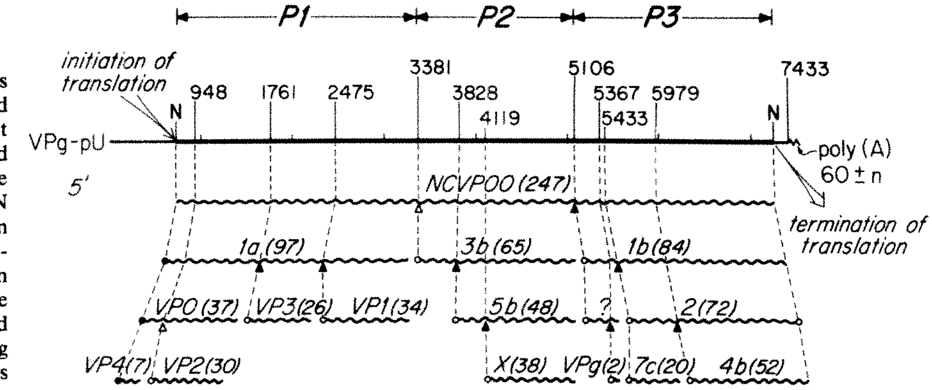


Fig. 4 Gene organization of poliovirus RNA. Virion RNA, terminated at the 5′ end with the genome-linked protein VPg and at the 3′ end with poly(A), is shown as a solid line, the translated region being more pronounced than the non-coding regions. N indicates the proposed sites at which initiation (N741) and termination (N7361) of translation occurs. The numbers above the virion RNA line refer to the first nucleotide of the codon specifying the N-terminal amino acid for the viral specific proteins. The coding region has been divided into three regions (P1, P2, P3), corresponding to the primary cleavage products of polyprotein NCVPOO. Accordingly, polypeptides (waved line, designated in italics) are referred to in the text as P1-1a, P3-1b, etc. No 'leader peptide' that would map in a region preceding P1 has been found for poliovirus RNA translation. Numbers in parentheses are molecular weights calculated from the amino acid sequences shown in Fig. 2, assuming that C-terminal 'trimming' does not occur. Open circles indicate that the exact position of the terminal acid is known. The N-termini are glycine in all cases except for VP2 where it is serine. The C-terminal amino acid of P3-2 is phenylalanine. Closed circles indicate that the N-termini are known to be blocked. Closed triangles: Gln-Gly pairs that are cleaved during proteolytic processing of a polypeptide. Open triangles: exceptional cleavage sites in polypeptide processing (Asn-Ser at N948; Tyr-Gly at N3381). This figure is not meant to imply that the sequence of processing occurs via all possible intermediate polypeptides. For example, it is not known whether P2-3b is always processed to P2-X via P2-5b, or occasionally directly to P2-X.

shown but will be discussed in a later section. Omitted are also several small virus-specific peptides ($M_r < 12,000$) that can be observed in PAGE analyses of infected cells, whose origins and functions, however, are unknown.

The 3′-terminal, non-coding sequence

Our data suggest that the 3′-terminal, non-coding sequence of poliovirus RNA consists of a heteropolymeric segment of 72 nucleotides followed by poly(A). It lacks the polyadenylation signal A-A-U-A-A-A- found in polyadenylated mRNAs of eukaryotic cells[59]. Lack of this hexanucleotide supports our notion that the poly(A) of poliovirus RNA is genetically coded in that it is transcribed from 5′-terminal poly(U) in viral minus-stranded RNA[39,60-62]. Poly(A) in poliovirus RNA is very heterogeneous in length[18], but the average chain length has been determined directly by rapid sequencing techniques to be 60 residues long[63]. The function of poly(A) in poliovirus RNA is obscure, but it may have a role in the initiation of synthesis of minus-stranded RNA[64,65].

Signals in proteolytic processing

With a single exception (VP2) all viral polypeptides that we have analysed have glycine as their N-terminal amino acid (Fig. 2; refs 7, 36). Inspection of the sequences of the precursor polypeptides reveals that in 8 of 9 cases the NH_2-terminal Gly of the cleavage products was preceded by a glutamine residue. These data strongly suggest that proteolytic processing of poliovirus polypeptides occurs predominantly at Gln-Gly pairs in all three regions of the polyprotein, and that these cleavages are carried out by the same proteinase.

It has been assumed that host proteinase(s) act on the polyprotein NCVPOO to cut it *in statu nascendi* into the 'primary' cleavage products P1-1a, P2-X and P3-1b[67]. Proteolytic processing of P1-1a and P3-1b into 'secondary' cleavage products, on the other hand, has been clearly shown to be catalysed by virus-coded enzyme(s)[48,49,68]. We now consider it likely that the primary cleavage that generates P3-1b and secondary cleavages are carried out by the viral proteinase(s) because (1) the cleavage sites recognized in these processes are identical or similar (Fig. 2), and (2) a viral polypeptide can undergo autocatalytic cleavage to yield a viral proteinase and a putative RNA polymerase. The latter case exists in EMC virus, where polypeptide D (corresponding to P3-2 in polio) cleaves to proteinase p22 and E[69], corresponding to P3-7c and P3-4b, respectively, in poliovirus. Assuming that an analogous situation exists for poliovirus, it is tempting to speculate that during initial stages of infection primary cleavages result from autocatalytic events or by *trans* action of one polyprotein on another. As the infection proceeds, sufficient amounts of mature viral proteinase are produced to catalyse subsequently primary as well as secondary cleavages.

Inspection of the amino acid sequence illustrated in Fig. 2 shows the dipeptide Gln-Gly to occur 13 times in the polyprotein NCVPOO, 2.5 times the expected number of occurrences for this amino acid pair calculated assuming random distribution of amino acids of the composition and size of the polyprotein. The high Gln-Gly frequency supports the hypothesis that Gln-Gly is a major signal for processing of poliovirus proteins. At least 8 of the 13 Gln-Gly pairs are utilized in processing. Gln-Gly cleavage may also occur at N5796 to yield polypeptide P3-4a[70], and at N6486 to yield the products P3-6a and P3-6b of an alternate cleavage pathway within P3-1b[58]. Amino acid sequence analyses of these products are now in progress. Remaining Gln-Gly pairs at N1302, N3519 and N4626 may be ignored in processing either because the amino acid sequences surrounding these dipeptides may protect them from the proteinase, or because these cleavage signals may be sequestered. Alternatively, these Gln-Gly pairs may be utilized to a lesser degree so that the corresponding cleavage products have not been detected. Preliminary studies of the amino acid sequences around the Gln-Gly pairs that are known processing sites have not revealed other constant features likely to contribute to processing specificity. Other processing determinants besides the Gln-Gly sequence must exist, however, because infection with poliovirus does not result in detectable degradation of host cellular proteins, at least not for the first 4 h during which the bulk of viral macromolecular synthesis and proteolytic processing occurs[67].

One amino acid pair of interest is the Asn-Ser in VPO (at N945) that is cleaved during maturation of the virion to yield VP4 and VP2 (reviewed in ref. 14). Asn-Ser (two helix-breaker residues) differs from Gln-Gly (helix maker–helix breaker residues)[71] and could therefore be part of a signal for a processing step at maturation catalysed by a unique proteinase. A possible candidate, P2-X, has been reported to have proteolytic activity[72]. Note that the Asn-Ser dipeptide is predicted to occur at seven other locations in the polyprotein, all but two of which are within the capsid region.

The cleavage that generates the N-terminus of P2-3b is predicted by our sequence (Fig. 2) to occur between a tyrosine and a glycine residue. We have confirmed that the N-terminus of P2-3b is glycine (unpublished results); however, the identity of the C-terminus of VP1 has not yet been directly determined. Thus we cannot yet be certain whether this site is unique and is cleaved by another, perhaps cellular, proteinase, whether it is cleaved by the activity that recognizes Gln-Gly pairs, or whether a sequencing error has resulted in the erroneous prediction of tyrosine for this position.

Previous studies have shown that a Gln-Gly may be cleaved between the VP3 and VP1 of mengoviruses[73]. In contrast to our data, the NH_2-terminal amino acids of mengovirus are Asp in VP2, Ser in VP3, and Gly in VP1[73], and of aphthovirus are Asp in VP2, Gly in VP3, and Thr in VP1[76]. Burrel and Cooper[74] reported Gly, Asp and Ser as N-terminal amino acids in a mixture of poliovirus capsid proteins. Our nucleotide sequence predicts at the N-termini Ser in VP2 and Gly in VP3 and VP1, which we have confirmed by radiochemical microsequence analysis (Fig. 2). Thus, a discrepancy exists that cannot be explained at the present time.

Conclusions

To our knowledge, this work is the first complete description of the primary structure of a eukaryotic, helper-independent RNA virus genome. The protein-linked genome represents a unique chemical entity not previously found among naturally occurring RNAs.

The nucleotide sequence of the viral RNA is useful in several ways. In combination with radiochemical microsequence analyses we have determined the coding regions for each major viral polypeptide and thus their amino acid sequences. Based on the known virion structure[14] and the information presented in this article, the chemical structure of the poliovirion (calculated molecular weight 8.25×10^6) has been elucidated. Knowledge of the amino acid sequence will aid in the isolation and characterization of the viral gene products, for example, by antisera elicited by specific synthetic oligopeptides. The sequence data have also been instrumental in detecting and mapping cloned fragments of poliovirus cDNA. Specifically, the data have been useful in the prediction of cleavage sites of restriction enzymes and in the screening of clones by hybridization of $5'$-^{32}P-labelled oligonucleotides (Fig. 1) to the DNA in lysates of plasmid-containing *E. coli*[75].

The development of new vaccines by novel methods is desirable, and one approach is the construction of attenuated viral strains incapable of reverting to virulence[24]. A second approach could make use of the expression of cloned virus-specific genome segments in *E. coli* to produce viral antigens that are capable of eliciting formation of poliovirus neutralizing antibodies in man. Pursuit of both approaches requires that the genome structure and the gene organization of poliovirus be known.

Finally, construction of specific inhibitors of the viral proteinase(s) utilizing a knowledge of their cleavage specificity may lead to a novel set of antiviral chemotherapeutic agents[67].

We thank Akiko Kitamura, Lydia Fedus and Medora Hardy for technical assistance. We are indebted to Astrid Wimmer for support, to Carol Carter, Ann Jacobson, Nancy Reich and Akio Nomoto for helpful discussions, and to Keith Thompson for assistance in computing data. We also thank Joyce Schirmer and Jeffrey Demian for art work. Computer services were provided by the Stanford MOLGEN project of the NIH SUMEX-AIM facility. This work was supported in part by PHS grants AI 15122 and CA 28146; and by the US Department of Energy. B.L.S. is supported by NIH fellowship AI 05935; E.A.E. is a postdoctoral trainee; and C.J.A. a predoctoral trainee of NCI, grant 09176. E.W. is a Josiah Macy Jr Faculty Scholar.

Received 16 February; accepted 8 April 1981.

1. Landsteiner, K. & Popper, E. *Wien. Klin. Wochenschr.* **21,** 1830 (1908).
2. Salk, J. & Salk, D. *Science* **195,** 834–847 (1977).
3. Sabin, A. B. & Boulger, C. R. *J. biol. Standard* **1,** 115–118 (1973).
4. Lee, Y. F., Nomoto, A. & Wimmer, E. *Prog. Nucleic Acid Res. molec. Biol.* **19,** 89–95 (1976).
5. Lee, Y. F., Nomoto, A., Detjen, B. M. & Wimmer, E. *Proc. natn. Acad. Sci. U.S.A.* **74,** 59–63 (1977).
6. Nomoto, A., Detjen, B., Pozzatti, R. & Wimmer, E. *Nature* **268,** 208–213 (1977).
7. Semler, B. L. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
8. Larsen, G. R., Semler, B. L. & Wimmer, E. *J. Virol.* **37,** 328–335 (1981).
9. Pettersson, R. F., Flanegan, J. B., Rose, J. K. & Baltimore, D. *Nature* **268,** 270–272 (1977).
10. Nomoto, A., Kitamura, N., Golini, F. & Wimmer, E. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5345–5349 (1977).
11. Jacobson, M. F. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **61,** 77–84 (1968).
12. Summers, D. F. & Maizel, J. V. *Proc. natn. Acad. Sci. U.S.A.* **59,** 966–971 (1968).
13. Holland, J. J. & Kiehn, E. D. *Proc. natn. Acad. Sci. U.S.A.* **60,** 1015–1022 (1968).
14. Rueckert, R. R. *Compreh. Virol.* **6,** 131–213 (1976).
15. Cooper, P. D. *Compreh. Virol.* **9,** 133–207 (1977).
16. Kates, J. *Cold Spring Harb. Symp. quant. Biol.* **35,** 743–752 (1970).
17. Armstrong, J. A., Edmonds, M., Nakazato, H., Philips, B. A. & Vaughan, M. H. *Science* **176,** 526–528 (1972).
18. Yogo, Y. & Wimmer, E. *Proc. natn. Acad. Sci. U.S.A.* **69,** 1877–1882 (1972).
19. Kacian, D. L. & Myers, J. C. *Proc. natn. Acad. Sci. U.S.A.* **73,** 2191–2195 (1976).
20. Kitamura, N. & Wimmer, E. *Proc. natn. Acad. Sci. U.S.A.* **77,** 3196–3200 (1980).
21. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5463–5467 (1977).
22. Sanger, F. & Coulson, A. R. *FEBS Lett.* **87,** 107–116 (1978).
23. Nomoto, A., Jacobson, A., Lee, Y. F., Dunn, J. J. & Wimmer, E. *J. molec. Biol.* **128,** 179–196 (1979).
24. Nomoto, A. *et al. Virology* (in the press).
25. Kramer, F. R. & Mills, D. R. *Proc. natn. Acad. Sci. U.S.A.* **75,** 5334–5338 (1978).
26. Flanegan, J. B., Pettersson, R. F., Ambros, V., Hewlett, J. J. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **74,** 961–965 (1977).
27. Hewlett, M. J. & Florkiewicz, R. Z. *Proc. natn. Acad. Sci. U.S.A.* **77,** 303–307 (1980).
28. Nomoto, A. & Imura, N. *Nucleic Acids Res.* **7,** 1233–1246 (1979).
29. Nomoto, A. *et al. J. molec. Biol.* **128,** 165–177 (1979).
30. Harris, T. J. R., Dunn, J. J. & Wimmer, E. *Nucleic Acids Res.* **5,** 4039–4054 (1978).
31. Nomoto, A., Lee, Y. F. & Wimmer, E. *Proc. natn. Acad. Sci. U.S.A.* **73,** 375–380 (1976).
32. Rothberg, P. G., Harris, T. J. R., Nomoto, A. & Wimmer, E. *Proc. natn. Acad. Sci. U.S.A.* **75,** 4868–4872 (1978).
33. Ambros, V. & Baltimore, D. *J. biol. Chem.* **253,** 5263–5266 (1978).
34. Wu, M., Davidson, N. & Wimmer, E. *Nucleic Acids Res.* **5,** 4711–4723 (1978).
35. Golini, F., Nomoto, A. & Wimmer, E. *Virology* **89,** 112–118 (1978).
36. Kitamura, N. *et al. Cell* **21,** 295–302 (1980).
37. Pallansch, M. A. *et al. J. Virol.* **35,** 414–419 (1980).
38. Holzer, H. & Wohlhueter, R. *Adv. Enzyme Reg.* **10,** 121–132 (1972).
39. Yogo, Y. & Wimmer, E. *J. molec. Biol.* **92,** 467–477 (1975).
40. Pettersson, R. F., Ambros, V. & Baltimore, D. *J. Virol.* **27,** 357–365 (1978).
41. Wimmer, E. in *The Molecular Biology of Picornaviruses* (ed. Pérez-Bercoff, R.) 175–190 (Plenum, New York, 1979).
42. Kitamura, N., Adler, C. & Wimmer, E. *Ann. N.Y. Academy Sci.* **354,** 183–201 (1980).
43. Ehrenfeld, E. in *The Molecular Biology of Picornaviruses* (ed. Pérez-Bercoff, R.) 223–238 (Plenum, New York, 1979).
44. Steitz, J. A. in *RNA Phages* (ed. Zinder, N. D.) 319–352 (Cold Spring Harbor Laboratory, New York, 1975).
45. Golini, F., Semler, B. L., Dorner, A. J. & Wimmer, E. *Nature* **287,** 600–603 (1980).
46. Shine, J. & Dalgarno, L. *Proc. natn. Acad. Sci. U.S.A.* **71,** 1342–1346 (1974).
47. Kozak, M. in *Protein Biosynthesis in Eukaryotes* (Plenum, New York, in the press).
48. Palmenberg, A. C., Pallansch, M. A. & Rueckert, R. R. *J. Virol.* **32,** 770–778 (1979).
49. Svitkin, Yu. V., Gorbalenya, A. E., Kazachkov, Yu. A. & Agol, V. I. *FEBS Lett.* **108,** 6–9 (1979).
50. Lundquist, R. E., Ehrenfeld, E. & Maizel, J. V. Jr *Proc. natn. Acad. Sci. U.S.A.* **71,** 4773–4777 (1974).
51. Flanegan, J. B. & Baltimore, D. *J. Virol.* **29,** 352–360 (1979).
52. Wouters, M. & Vandekerckhove, J. *J. gen. Virol.* **33,** 529–533 (1976).
53. Wetz, K. & Habermehl, K. O. *J. gen. Virol.* **44,** 525–534 (1979).
54. Anderson, C. W. & Lewis, J. B. *Virology* **104,** 27–41 (1980).
55. Sanger, D. V., Black, D. N., Rowlands, D. J., Harris, T. J. R. & Brown, F. *J. Virol.* **33,** 59–68 (1980).
56. Vrijsen, R., Boéye, A. & Strosberg, A. D. *Biochem. biophys. Res. Commun.* **85,** 1596–1601 (1978).
57. Heukeshoven, J. & Dernick, R. *Z. Naturforschung* **36,** 164–172 (1981).
58. Rueckert, R. R., Matthews, T. J., Kew, O. M., Pallansch, M., McLean, C. & Omilianowski, D. in *The Molecular Biology of Picornaviruses* (ed. Pérez-Bercoff, R.) 113–125 (Plenum, New York, 1979).
59. Fitzgerald, M. & Shenk, T. *Cell* **24,** 251–260 (1981).
60. Yogo, Y., Teng, M. H. & Wimmer, E. *Biochem. biophys. Res. Commun.* **61,** 1101–1109 (1974).
61. Spector, D. H. & Baltimore, D. *Virology* **67,** 498–505 (1975).
62. Dorsch-Häsler, K., Yogo, Y. & Wimmer, E. *J. Virol.* **16,** 1512–1527 (1975).
63. Ahlquist, P. & Kaesberg, P. *Nucleic Acids Res.* **7,** 1195–1204 (1979).
64. Spector, D. H. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **71,** 2983–2987 (1974).
65. Dasgupta, A., Zabel, P. & Baltimore, D. *Cell* **19,** 423–429 (1980).
66. Porter, A. G. *et al. Nature* **276,** 298–301 (1978).
67. Korant, B. D., Chow, N., Lively, M. & Powers, J. *Ann. N.Y. Acad. Sci.* **343,** 304–315 (1980).
68. Pelham, H. R. B. *Eur. J. Biochem.* **85,** 457–462 (1978).
69. Rueckert, R. R., Palmenberg, A. C. & Pallansch, M. A. *Eur. J. cell. Biol.* **22,** 263–275 (1980).
70. Etchison, D. & Ehrenfeld, E. *Virology* **107,** 135–143 (1980).
71. Anfinsen, C. B. & Scheraga, H. H. *Adv. Protein Chem.* **25,** 205–300 (1975).
72. Korant, B., Chow, N., Lively, M. & Powers, J. *Proc. natn. Acad. Sci. U.S.A.* **76,** 2992–2995 (1979).
73. Scraba, D. G. in *The Molecular Biology of Picornaviruses* (ed. Pérez-Bercoff, R.) 1–23 (Plenum, New York, 1979).
74. Burrel, C. J. & Cooper, P. D. *J. gen. Virol.* **21,** 443–451 (1973).
75. Van der Werf, S. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
76. Boothroyd, J. C. *et al. Nature* **290,** 800–802 (1981).

Effects of sleep and arousal on the processing of visual information in the cat

Margaret S. Livingstone[*] & David H. Hubel

Department of Neurobiology, Harvard Medical School, Boston, Massachusetts 02115, USA

Single units in the cat lateral geniculate nucleus and primary visual cortex show changes in both spontaneous and visually evoked firing as a function of the state of wakefulness. On arousal spontaneous firing is smoother and often reduced, whereas evoked responses are usually enhanced; the result is an increase in the signal-to-noise ratio. Single- and double-label 2-deoxyglucose autoradiographs show further that slow-wave sleep differentially depresses visually evoked activity in the deeper layers of the visual cortex.

SUBJECTIVELY it seems obvious that the way in which the brain handles sensory information must change radically between sleeping and waking. To examine these changes it is important to be able to influence the firing of neurones along a sensory pathway by natural stimulation. We have chosen the visual system to study neural activity as a function of the state of arousal because, in the pathway from the retina through the lateral geniculate body (LGB) to the primary visual cortex, the properties of the cells are well understood in terms of their optimum visual stimuli. In an animal that closes its eyes when asleep, additional mechanisms for blocking visual messages during sleep might seem unnecessary but, beyond the retina, such mechanisms do seem to exist. In the cat, spontaneous firing patterns of cells in the LGB[1,2] and primary visual cortex[3-6] show marked differences between sleeping and waking. In the LGB Coenen and Vendrik[7] have shown that responses to light stimulation are strongly enhanced on arousal in lightly anaesthetized cats. In the present study we have recorded from single cells in the LGB and striate cortex[8], examining the effects of the state of wakefulness on both the vigour and specificity of responses. We have also compared sleep and arousal using the 2-deoxyglucose method for labelling neurones according to activity[9,10].

Single-cell recording

We compared spontaneous and visually evoked firing in sleeping and waking states in 130 cortical cells in 15 cats, and in 14 geniculate neurones in 2 cats.

Each cat was first sleep deprived overnight in a slowly revolving (0.5 r.p.m.) drum, and then prepared as if for chronic recording[1,4], under halothane anaesthesia. Four chlorided silver electroencephalogram (EEG) electrodes were cemented over the dura and a pedestal for holding a hydraulic microelectrode advancer was cemented to the skull. For cortical recording we positioned the pedestal over the postlateral gyrus. For geniculate recording the animal's head was put in a Horsley–Clarke headholder and the pedestal cemented stereotaxically above the LGB. The head was then freed, well before terminating the anaesthesia. We infiltrated skin incisions with Xylocaine, intubated the trachea, paralysed the animal with continuous intravenous succinylcholine and monitored levels of expired CO_2. The head was held firmly with adhesive tape and gauze pads. Pupils were dilated with homatropine, contact lenses were fitted so that images on a screen at 1.5 m distance produced a focused image on the retina, and the optic disk and area centralis of each eye were mapped by optical projection. When necessary we used an adjustable prism to bring the two centres of gaze into superposition.

The halothane was then discontinued, and in the EEG irregular high-voltage activity of the anaesthetized state gave way, in a few minutes, to the low-voltage rapid activity of the waking state. Most of the cats then slept spontaneously and frequently, as indicated by high-voltage slow waves. Arousal from slow-wave sleep was spontaneous or could be produced by a noise or tactile stimulus. Rapid eye movement (REM) sleep was recognized by episodes of large monophasic potentials recorded from the visual cortex (ponto-geniculo-occipital (PGO) waves), which could be suppressed by arousing the animal.

For each cell we first mapped receptive fields on the screen with a hand-held slide projector. For cortical cells a computer-driven optic bench was adjusted to produce moving-slit stimuli which were tailored for optimal position, orientation, shape and rate of movement. The stimuli for geniculate cells were stationary circular spots or annuli. Action potentials, EEGs and traces representing the presentations of the stimuli were recorded on a Grass polygraph and on magnetic tape for subsequent averaging. Every cell included in the series was followed over several transitions between waking and high-voltage slow-wave sleep.

Recordings from lateral geniculate body

With the screen diffusely lit, geniculate cells showed an increase in spontaneous firing rate when the animal was aroused from slow-wave sleep, as expected from previous studies[7,11-13]. Moreover, the pattern of firing changed from high-frequency bursts of two to eight spikes to more regular firing[1,3,14].

In all 14 geniculate cells there were marked changes also in responsiveness to visual stimuli, as illustrated in Fig. 1. In Fig. 1*a*, during an otherwise wakeful period, there was a brief episode of high-voltage slow-wave sleep lasting about 30 s and ending spontaneously. The stimulus, which was repeated throughout, was a small (1/4°) spot of light covering the receptive field centre. During the slow waves there was a depression of the on-response, obvious in both the spike record and the average-response histograms. This change is most easily explained by supposing that the efficacy of optic nerve fibres in triggering the geniculate cell is enhanced on arousal[7]. All the effects of arousal cannot be explained in this way, however. In Fig. 1*b*, the stimulus was an annulus covering the surround of the cell's receptive field. Now spontaneous arousal was accompanied not only by an increase in the off-discharge, but also by an enhanced suppression of firing while the annulus was on.

Figure 1*c* shows the enhancement of suppression for the same on-centre cell, in a slightly different way. In Fig. 1*c*, upper half, the screen was lit steadily and diffusely by a low-level background light. The tonic firing of the cell was obviously decreased during each of the two bursts of slow waves. In the lower half, an annulus covering the receptive field surround tonically sup-

---

[*] Present address: Department of Biology, Princeton University, Princeton, New Jersey 08544, USA.

0028-0836/81/250554—08$01.00

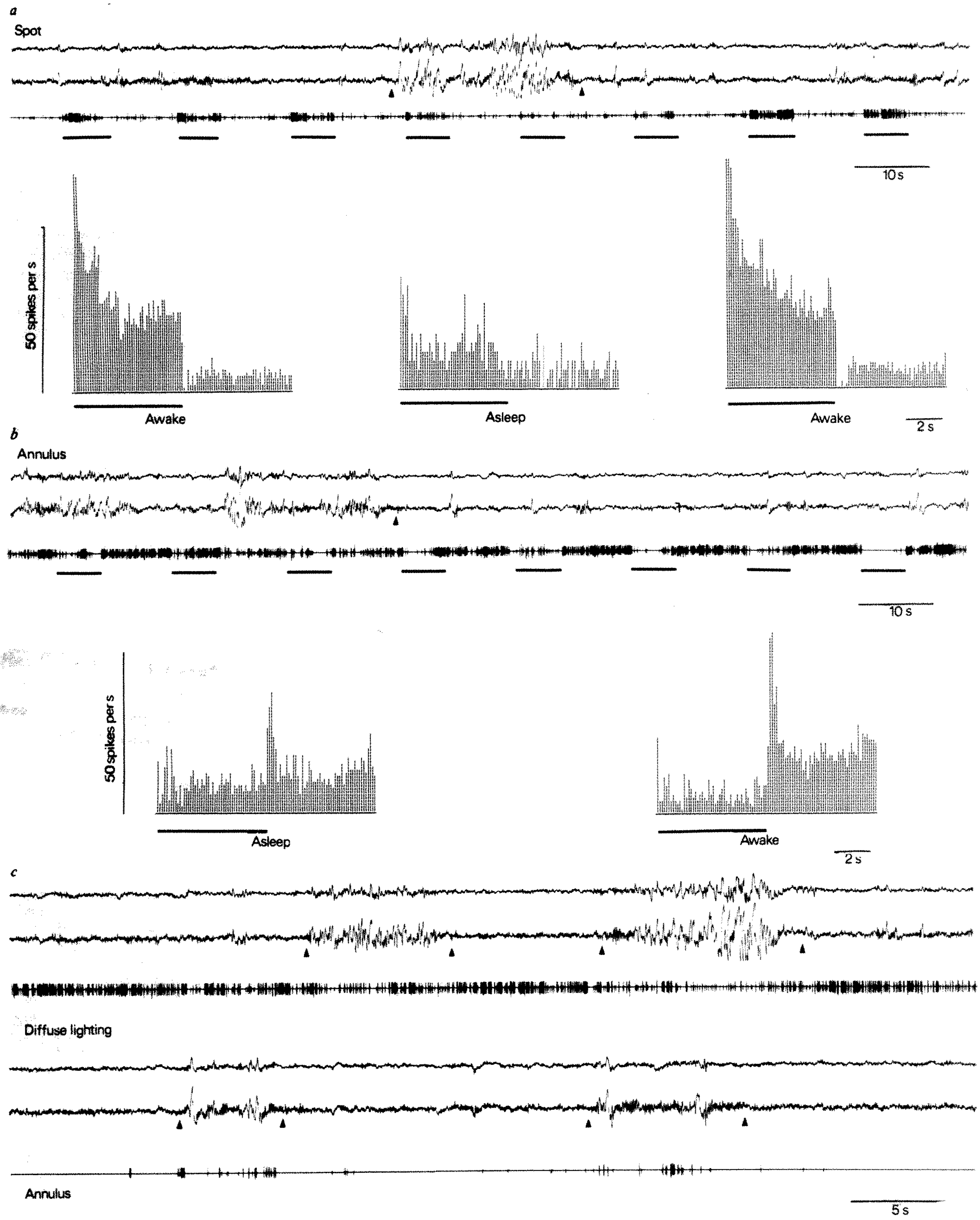


Fig. 1 *a* and *b*, Responses of an on-centre cell from layer A of the lateral geniculate nucleus in a drowsy cat. In each polygraph record the upper two traces are the bipolar EEGs from anterior and posterior head regions; the third trace shows the spikes from the geniculate cell, converted into pulses, in these and subsequent polygraph records, by a Schmitt trigger circuit. Each record lasts ~2 min. Brief periods of sleep, which end spontaneously, are indicated by the slow waves in the middle of record *a* and (roughly) the left half of *b*. The approximate transitions between the flat waking EEG record and periods of slow-wave sleep are indicated by arrowheads. In *a* a spot covering the receptive field centre is flashed on and off. The relative weakness of the on responses during periods of slow waves can be seen in the record of spikes and in the three histograms. The left histogram is the average of 13 responses, of which the last 3 are shown. The middle histogram is an average of the 2 responses during the burst of slow waves, and the right histogram the average of the next 16 responses. In *b* the stimulus is an annulus covering the receptive-field surround. Both the suppression of firing during the stimulus and the off discharge are weaker during sleep. Histograms are the average of nine responses each, before and after the arrowhead. *c*, Effects of slow-wave sleep on steady-state firing of the same cell. In the upper set of three traces (50 s) the screen is diffusely lit by the background light. At each burst of slow waves in the cortical EEG (upper two traces) the rate of firing declines. In the lower set an annulus covers the receptive-field surround. In the aroused periods the discharges are almost completely suppressed, but each episode of slow waves is accompanied by bursts of impulses. Onset and end of slow-wave bursts are indicated by arrowheads.

pressed the firing. The suppression was almost complete when the EEG was flat, but during the slow waves there were bursts of impulses. Thus arousal could depress or elevate the cell's firing rate, depending on the visual stimulus used.

For off-centre cells there was a similar enhancement, both in the suppression of firing and off-discharge to a spot covering the receptive field centre, and in the on-response to an annulus. In both on-centre and off-centre cells diffuse light was less effective than a spot confined to the receptive field centre. This is characteristic of geniculate cells. The diffuse light responses of both on-centre and off-centre cells showed little or no enhancement on arousal from slow-wave sleep, presumably because the increase in the response from the field centre was counterbalanced by an increased opponent response from the surround. To understand the significance of this one must recall that a weaker response to a large spot than to a small one just covering the field centre is characteristic also of retinal ganglion cells[15,16]. But geniculate cells receiving their main excitatory optic input from a single retinal ganglion cell show a more pronounced reduction in response to stimuli including the field surround than do the retinal ganglion cells[14,17]: indeed, this is the main known contribution of the lateral geniculate to the processing of visual information. If the response enhancement on arousal were greater for a large spot than for a small one, one would have to conclude that arousal led to an impairment of this function.

Arousal from slow-wave sleep produced no obvious changes in geniculate receptive field sizes as mapped by small spots. As visual acuity is presumably related in geniculate cells to receptive field centre size, an increase in centre size on arousal[18] would have been perplexing if not dismaying.

In summary, arousal from slow-wave sleep increased the spontaneous firing rate of lateral geniculate cells and reduced or eliminated the high-frequency bursts (not resolved in the slow time scale of Fig. 1); both the excitatory and inhibitory components of responses to spatially restricted visual stimuli were enhanced; and, finally, responses from the receptive field centre and the antagonistic surround were both enhanced.

Presumably the mechanism of these effects depends on nonretinal projections to the LGB because in mammals there do not seem to be projections from brain to retina, and firing of retinal ganglion cells is unaffected by arousal[1]. The increase in the on- or off-discharges on arousal is known to reflect an increase in the proportion of optic nerve impulses that succeed in triggering geniculate impulses (the firing index)[7]. The enhanced suppression of discharge seen on arousal may reflect enhanced inhibition from other geniculate neurones that are excited by the stimulus and whose firing index is increased by arousal; or arousal may lead to an enhancement of both excitatory and inhibitory synapses. An increased firing during sleep might also be somehow related to the high-voltage cortical slow waves, which, if present in thalamus and cortex concurrently, could be associated with swings in membrane potential capable of triggering bursts of impulses. Suppressing these slow waves, on arousal, could then abolish the discharges.

Visual cortex

In contrast to geniculate cells, cortical cells varied greatly from one to the next in the degree to which they were influenced by changes in arousal level. Marked differences were often seen between two cells simultaneously recorded from two electrodes spaced 1 mm or less apart, or between two cells simultaneously recorded from one electrode. Nevertheless, each individual cell showed a consistent type of change over several cycles of

[illegible] SCIENCE COLLEGE [illegible] Upper Circular Road [illegible]

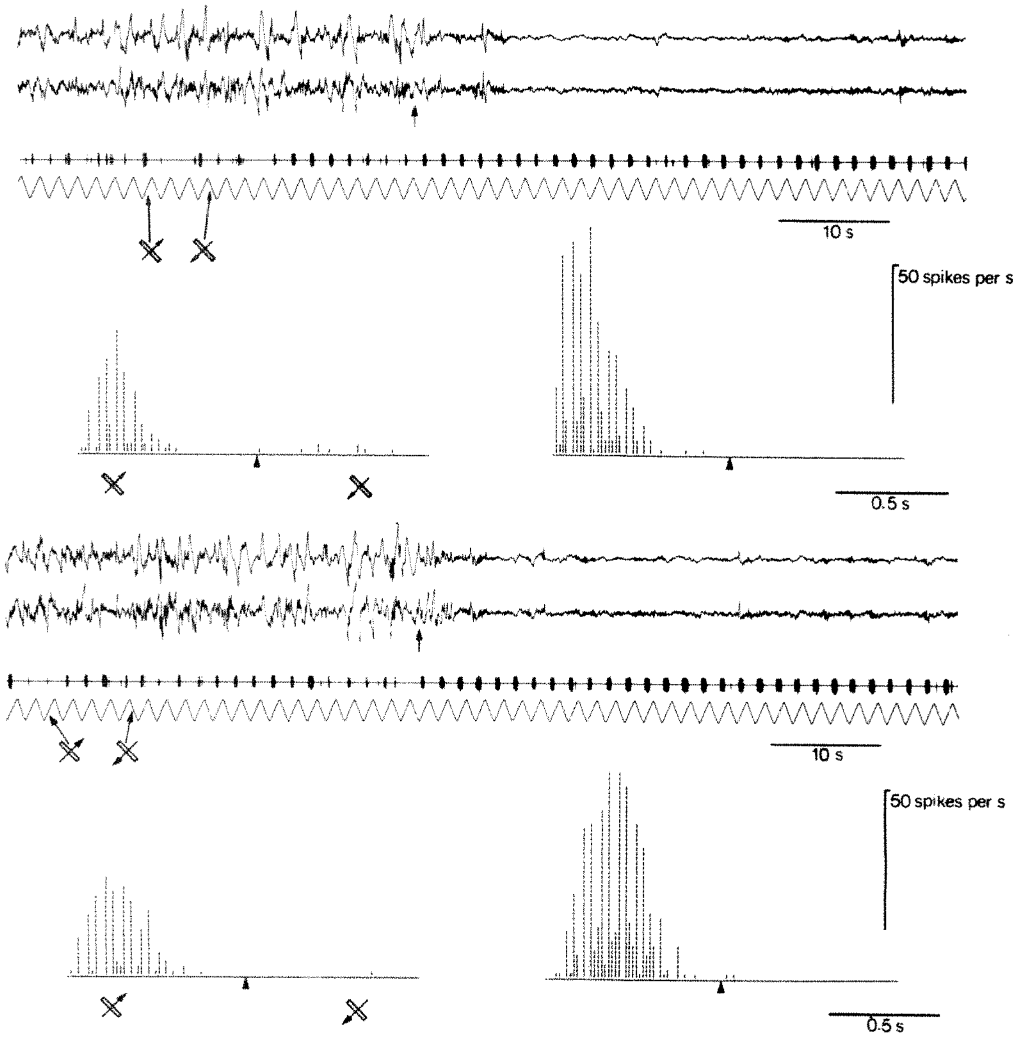


Fig. 2 Effects of arousal on the responses of a cell in layer 6 of striate cortex. The record is 85 s long. Near the middle a noise (arrow) arouses the cat, as shown by the suppression of the EEG slow waves in the upper pairs of traces. Responses of the cell (third trace) are to a 1/2°×5° stimulus oriented 45° anticlockwise to vertical and swept up and to the right and down to the left (indicated in the fourth trace by up-and-down deflections). On arousal the responses are more consistent and on the whole stronger, as can be seen in the average response histograms. Each histogram is the average of 40 responses. The arrowhead below each histogram indicates when the stimulus reversed direction.

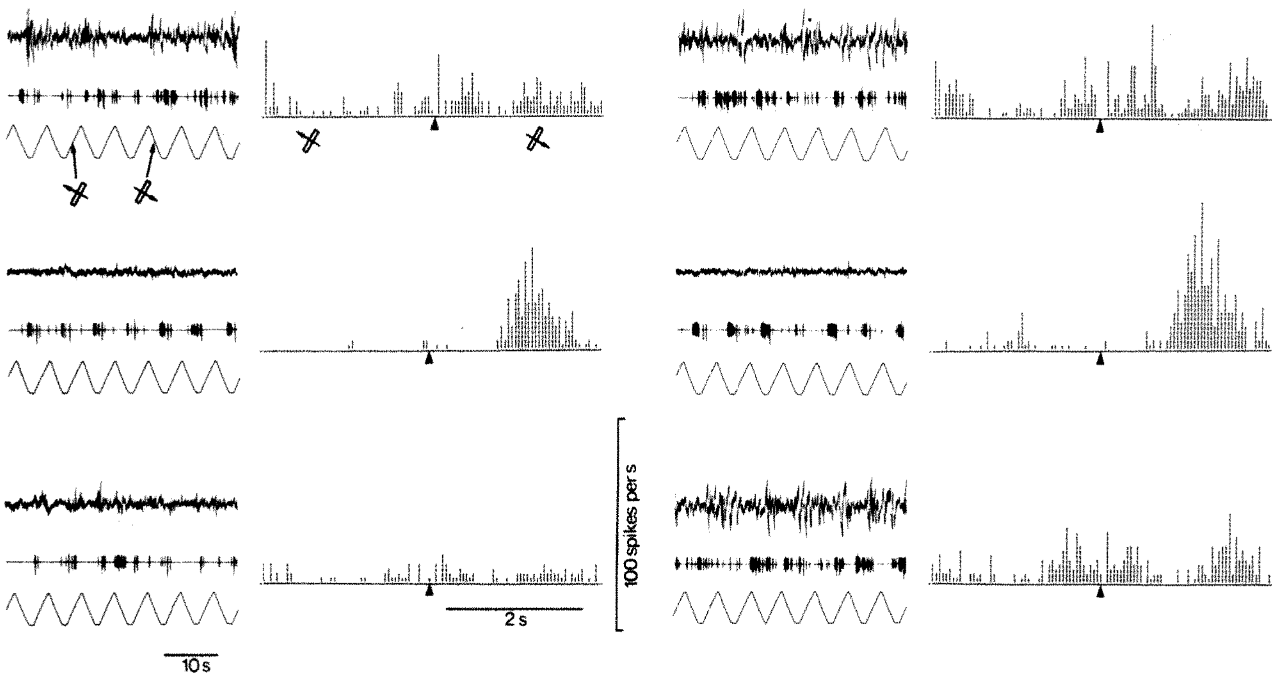


Fig. 3 Responses of a cell in the striate cortex with the cat in various levels of arousal. Each record lasts 45 s and ~3 min elapse between records. As indicated by the EEG (upper trace), the three records on the left side are taken, from above downwards, in slow-wave sleep, awake and slow-wave sleep; those on the right are in slow-wave sleep, awake and slow-wave sleep. The middle trace shows impulse discharges, and lower trace indicates by up-and-down deflections left and right movement of a 1/4° × 2° slit oriented 30° clockwise to vertical. In the alert state the cell responds briskly to rightward movement, but hardly at all to leftward movement. In drowsiness and sleep spontaneous firing is greatly increased, and responses to the moving slit are almost completely absent. Each histogram is the average of 12 responses.

sleeping and waking. Usually (in two-thirds of the 130 cells studied) an irregular burst of spontaneous firing in slow-wave sleep was reduced in overall rate and replaced by a smoother, more regular pattern on awakening. In 10 cells there was an increased firing rate and a smoothing. The rest of the cells showed no difference in spontaneous activity between sleeping and waking. Visually evoked responses were either unaffected or enhanced by awakening: in no cells were responses diminished. Roughly one-third of all cells showed an increase in evoked response that was obvious on the polygraph record; many more cells showed an enhancement of responses that were detected only by histogram averaging. Even in cells showing no increase in evoked response, the response was usually easier to observe against the more regular or reduced background firing.

Figure 2 shows an example of the effects of arousal on the responses of a cortical cell. The stimulus was an optimally oriented moving slit of light. The responses, which were weak and capricious during slow-wave sleep, became stronger and more consistent after arousal by a noise. Despite the increase in response to the optimum direction of movement (up and to the right), the reverse direction remained ineffective.

Figure 3 shows six extracts from a record of another cell, taken over a 15-min period during which the cat repeatedly slept and awoke. The visually evoked responses were drastically changed, from vigorous while awake to almost completely absent during sleep. The increase in the response in the awake state was especially obvious because the background firing was reduced.

Four of the cells showed a transient burst of smooth steady firing lasting for about 30 s when the animal awoke spontaneously or was aroused by a noise or touch. After the burst the firing remained regular, usually stabilizing at a rate about the same as that during sleep or slightly higher. Had these cells not been clearly and specifically driven by particular visual stimuli, they might have been classed as multimodal, given their tendency to respond to any stimulus capable of arousing the animal. We saw no cells in striate cortex that responded to sense modalities other than visual, except in association with arousal.

A fall in both orientation specificity and movement–direction selectivity has been reported following stimulation of the menencephalic reticular formation[19]. Our preliminary results with natural arousal do not, however, point to a loss of specificity; on the contrary there may sometimes be an improvement. Figure 4 shows that the response to an optimally oriented slit was clearly, though not dramatically, improved after arousal, caused in this case by a loud noise, while the background firing was reduced and smoother. Directional selectivity was slightly improved, in that the response to the less effective direction of movement declined. For most directionally selective cells, however, the responses to the two opposite directions changed roughly in proportion. If, as in Fig. 2, one direction of movement failed to evoke a response in sleep, the same was so after arousal. No cell lost directional selectivity after arousal from sleep.

In testing orientation selectivity with a hand-held visual stimulator, we usually found no change. A few cells showed a slight improvement, but to be certain of this it will be necessary to compare orientation tuning curves.

Finally, in several cells we studied the effects of awakening on end stopping—the decline in response that occurs in some cells, termed hypercomplex, when a slit, edge, or dark bar is made longer than optimal[20]; the degree of stopping did not vary with level of arousal. The enhanced excitation from the activating part of the receptive field is therefore presumably offset by an enhanced inhibition from the outlying areas.

These observations are consistent with previous ones[20–22] that deepening barbiturate anaesthesia suppresses responses but does not change their specificity or lead to other qualitative changes. They also confirm the finding of Wurtz[23] that cells in area 17 recorded from awake alert monkeys with chronic recording techniques show responses that are not obviously different from responses obtained under barbiturate anaesthesia.

Most of the cells we studied were complex in type. The few simple cells examined showed the same types of changes as the complex cells. Though we have too few cells to make reliable comparisons of laminar distribution, the most striking changes in responses seemed to be in deeper-layer cells.

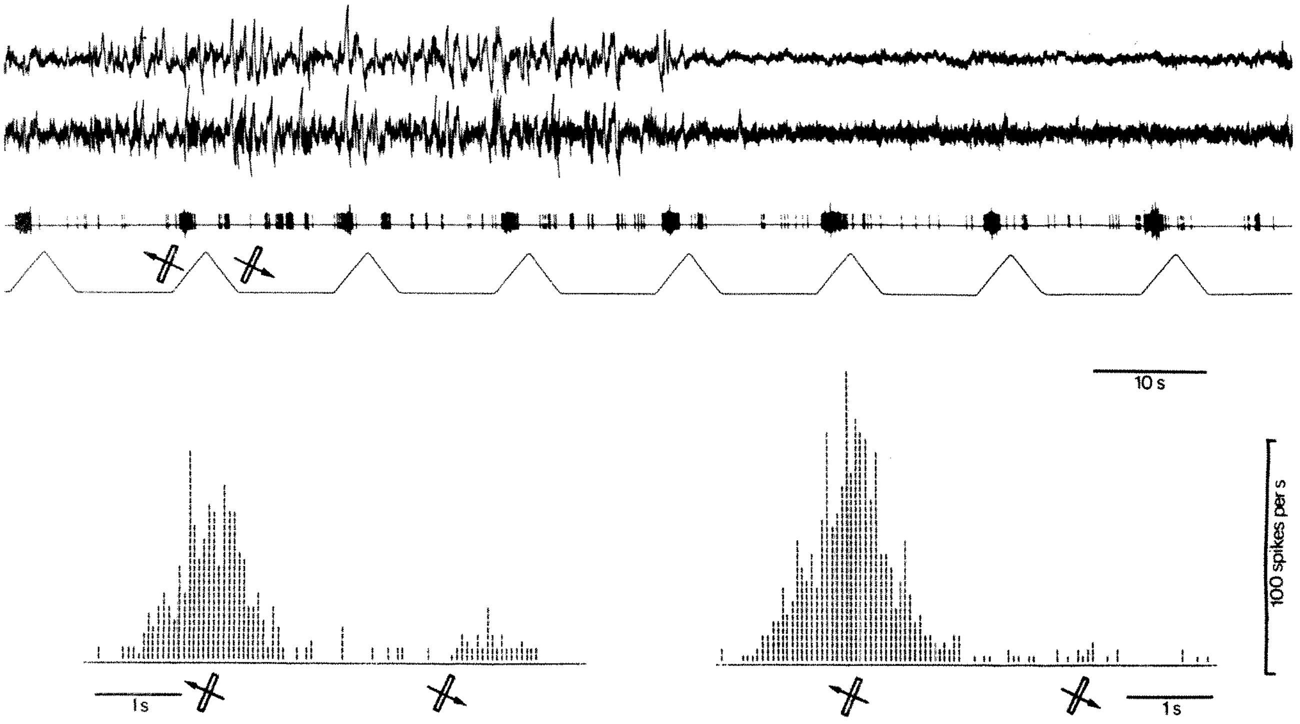


Fig. 4 Effects of arousal from slow-wave sleep on responses and response selectivity of a cell in layer 2 of striate cortex. About halfway through the 2-min record the cat is aroused by a noise. An optimal slit, 1/2° × 3°, oriented 25° clockwise to vertical, evokes a response (third trace) that is much greater to movement up and to the left than down and to the right. Arousal results in a moderate increase in the response to leftward movement, and a virtual elimination of the response to rightward movement (see the histograms). Arousal also produces suppression and smoothing of the spontaneous firing.

2-Deoxyglucose studies

We examined deoxyglucose uptake as a function of waking state in eight cats. Four cats, two awake and two in slow-wave sleep, were stimulated with moving vertical stripes confined to the right visual field. The other four cats were half the time awake and half the time asleep (see below). The cats studied in slow-wave sleep were sleep deprived on a treadmill.

Cats were prepared in the same way as those used for physiological recordings, except that the microelectrode advancer was not installed. After discontinuing the anaesthetic we waited until the arrhythmic slow-wave pattern of halothane anaesthesia gave way to the low-voltage rapid activity characteristic of arousal. In the sleep-deprived animals this waking activity was soon replaced by the high-voltage slow waves of slow-wave sleep. A noise or tactile stimulus was sufficient to arouse the animals.

The visual stimulus was a moving set of irregularly spaced vertical black and white stripes projected on a screen by a slide projector and a motor-driven mirror, which produced a back-and-forth movement at 2° s^{-1}. Stripe width was about 1/4° near the vertical midline, gradually increasing to 2–3° in the periphery. The stripes were confined to the right field of vision, and did not include the vertical midline but extended from 2–3° to the right of the midline, out to ~45°. The diffusely lit part of the screen (0.5 log cd m^{-2}) covered the vertical meridian and left visual field. We injected ^{14}C-2-deoxyglucose in a single pulse over 5–10 s, began the stimulus immediately and continued it for 45 min.

The cats were then deeply anaesthetized with thiopental and perfused with normal saline followed by formol–saline. The brains were removed, frozen in liquid N_2–cooled Freon-25 at −100 °C and sectioned on a cryostatat at −30 °C. Sections were mounted on glass coverslips, heated to 75 °C to drive off water and exposed against X-ray film. We sometimes combined autoradiographs of 2–4 successive cortical sections by making enlargements on film and superimposing them, so that we could average the patterns seen in several single sections and obtain an enhancement of the overall pattern.

Though the two waking brains differed consistently from the two sleeping ones, we were worried that the differences might be due to variability between individual cats rather than to differences in their arousal state. In our first attempt to demonstrate differences in a single animal we alternated the stimuli, presenting stripes to the left visual field when the cat was awake and to the right field when it was asleep. The result was unsatisfactory, perhaps because each hemisphere was stimulated only half the time and any label associated with visual stimuli would be seen against a background representing both states of wakefulness.

To overcome these problems (and also with an eye to countless other applications), we developed a double-label 2-deoxyglucose technique, exploiting the differential sensitivity to ^{14}C and ^{3}H β-emissions of standard X-ray film and LKB Ultrofilm ^{3}H. In standard X-ray film the emulsion is protected by a gelatin layer, ~1 μm thick, which acts as a partial barrier to the low-energy ^{3}H emissions. The LKB film lacks this coating and is roughly 20 times more sensitive than X-ray film to ^{3}H, whereas the two films are about equally sensitive to ^{14}C. These ratios were determined by exposing both types of film to standards made by soaking pieces of filter paper in serial dilutions of ^{3}H and ^{14}C. In the autoradiographs shown here we increased the ratio of sensitivity to ^{14}C to sensitivity to ^{3}H of the X-ray film to about 100:1 by interposing an additional 10-μm layer of polyethylene film.

In the injections of deoxyglucose we tried to achieve a $^{14}C/^{3}H$ ratio that would allow the ^{3}H to overpower the ^{14}C on the LKB film, and the ^{14}C to predominate on the X-ray film.

Both double-label cats were sleep deprived on a treadmill and prepared as described for the single-label (^{14}C) cats. After the halothane was discontinued we waited until the EEG had shifted to a normal slow-wave sleep pattern from which the animal could easily be aroused. We then injected the ^{14}C-2-deoxyglucose and stimulated the left visual field with vertical black and white stripes for 45 min, all the while praying that the animal would stay asleep. Next we aroused the cat, injected the ^{3}H-deoxyglucose and stimulated the right visual field for 45 min, keeping the cat awake by a general uproar and tactile stimuli as needed. To be certain of confining the stimulus to the contralateral hemisphere we avoided stimulating a strip of visual field 2–3° to either side of the vertical midline. The cat was perfused and the brain sectioned as already described. The histological sections (together with a set of filter paper standards) were exposed first to X-ray film protected by polyethylene film and then to LKB film.

Lateral geniculate autoradiography

In the two single-label animals whose right visual fields were stimulated when awake, the left LGBs were heavily labelled. The region of dense label was confined to the part of the geniculate corresponding to the area of visual field stimulated, extending on coronal sections medially almost to the medial border, which corresponds to the vertical meridian, and laterally over half to two-thirds of the geniculate's width, corresponding to the 45° lateral extent of the stimulus (Fig. 5*a*). The right (unstimulated) LGB showed label only at background levels. In one of the two animals stimulated while in slow-wave sleep, both sides showed label only at background levels; in the other animal there was only very faint labelling on the stimulated side (not shown). In the double-label cat (Fig. 5*b*) only the ^{3}H-sensitive film showed a left–right difference in geniculate labelling, and again the region with increased label (in the left geniculate) corresponded exactly to the part of the visual field stimulated. No difference was seen between the two sides on the polyethylene-protected X-ray film, except for slightly heightened labelling on the left, in exactly the same place as the label on the ^{3}H film and doubtless due either to some unfiltered ^{3}H or to some ^{14}C-deoxyglucose still in the cat's bloodstream 45 min after the first injection, when the second, waking, ^{3}H stimulus was begun. That the lack of ^{14}C label on the right side was not due to failure of uptake of ^{14}C is clear from the cortical autoradiographs in the same animal (see below). Considering the profound effects of slow-wave sleep on responses of the individual cells we studied, the pronounced differences in the autoradiography are not too surprising.

Cortical autoradiography

A consistent and marked difference was also seen in the pattern of labelling in the visual cortex. Figure 5*c* shows representative overlaid sections from two single-label (^{14}C) cats, awake and in slow-wave sleep. These animals both had had their right visual fields stimulated, and in each there were conspicuous columns in the left hemisphere, and no hint of any periodic labelling pattern in the right hemisphere. In the awake animal the columns extended through the full thickness of cortex (except for layer 1) with roughly uniform density. In the cats stimulated during slow-wave sleep, the columns in the upper layers (above layer 4) were labelled about as densely as in the waking cats, but below layer 4 they were in most sections hardly visible against the background. In some cats layer 4 was continuously and densely labelled (Fig. 5*c*, waking) whereas in other cats it was not (Fig. 5*c*, slow-wave sleep). These differences in layer 4 labelling did not consistently correlate with sleep state and we cannot account for them.

The double-label cat gave the same result (Fig. 6). The upper pair of autoradiograph overlays are made from the same sets of sections as the lower pair, the only differences being in the type of film used and the interposition of a sheet of polyethylene film in the lower (^{14}C) set. Columns extended through the full thickness of cortex in the waking, ^{3}H-labelled side, and for the most part only halfway down on the sleeping, ^{14}C side. The figure also demonstrates how effectively this technique can distinguish between two different stimuli, as in each case only the appropriate hemisphere shows columns.

Discussion

One might have imagined that when an animal awakens the neurones in its brain would simply become more active, as suggested by Sherrington[24]. While this may apply to some areas

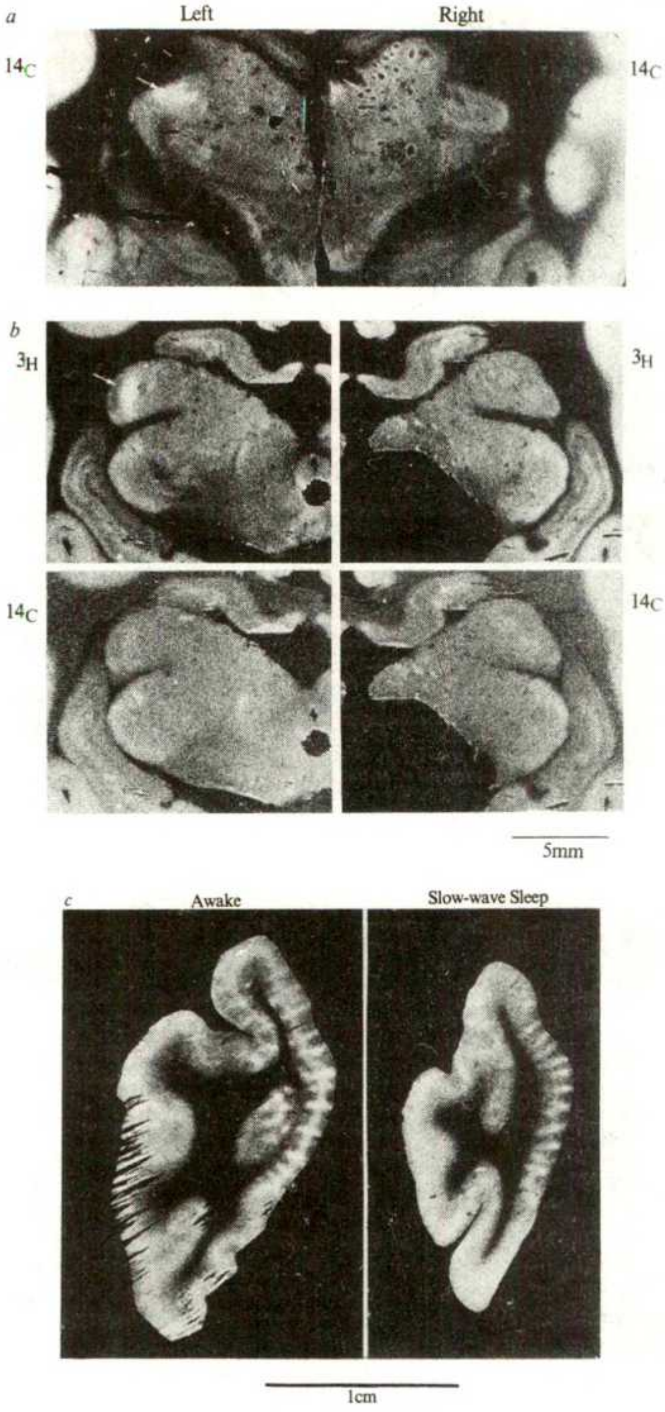


Fig. 5 *a*, 2-Deoxyglucose labelling of lateral geniculate nucleus in an awake cat by a moving grating of vertical black and white stripes. Stimulus was confined to the right field of vision, coming to within 2° or 3° of the vertical midline. Label in the left LGB is indicated by the arrow. Another waking cat gave a similar result. Figs 5 and 6 are negatives; labelled regions show as bright areas. *b*, Four autoradiographs showing results of the sleep–waking double-label experiment. After the ^{14}C-2-deoxyglucose injection (100 μCi kg^{-1}) the cat was stimulated while in slow-wave sleep for 45 min, the stimulus being confined to the left visual field. ^{3}H-2-deoxyglucose (5 mCi kg^{-1}) was then given, and followed by 45 min of stimulation to the right visual field during which the cat was kept aroused. The same sets of sections were exposed to LKB Ultrofilm (^{3}H) and to polyethylene-protected X-ray film (^{14}C). The ^{14}C autoradiographs of coronal sections show no obvious labelling; the ^{3}H autoradiographs by contrast show a dense patch of label in the appropriate part of the left LGB. *c*, Deoxyglucose labelling of visual cortex in two cats, awake and in slow-wave sleep, stimulated with vertical stripes confined to the right visual field. Each part consists of three superimposed successive sections. Columns are labelled on the medial surface (to the right) in both coronal sections through the left hemispheres; the vertical midline projection (postlateral gyrus, at the top) is, as expected, not labelled. In the waking cat the columns (which appear bright in these negatives) extend through the full cortical thickness; in the slow-wave sleep cat they are well labelled in the superficial half of the cortex, but below that they are faint or absent.

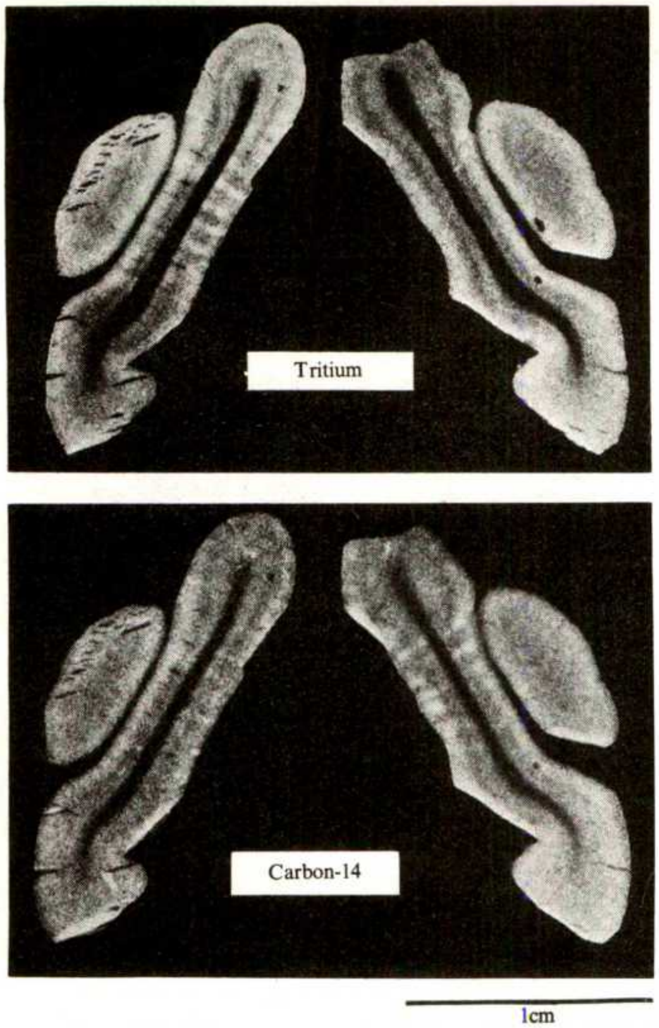


Fig. 6 Double-label deoxyglucose labelling of cortex; three successive sections exposed to LKB film (^{3}H) and to protected X-ray film (^{14}C) are superimposed. In the upper, waking, ^{3}H-labelled autoradiograph, columns are seen only in the left hemisphere and extend through the full cortical thickness. In the lower, sleeping, ^{14}C-labelled autoradiograph they are more densely labelled in the upper layers than in the lower ones.

of the brain, for example to parts of the brain stem reticular formation, it does not fit well with what we know of the cerebral cortex. The EEG does not elucidate this question, because although it is a sensitive indicator of arousal, its interpretation in terms of underlying neuronal activity is equivocal: a flattening on arousal from slow-wave sleep can be due to a reduction of activity or a desynchronization; the waves themselves may represent summation of impulses or of graded potentials. Recordings from single cortical neurones have shown that the flattening of the EEG on arousal from slow-wave sleep is most often associated with a lowered rate of spontaneous neuronal firing[4,5,25]. Some cells, in our experience a minority, show a transient increase in firing rate. Usually high-frequency bursts are replaced by more regular firing.

Early studies of the effects of arousal on responses from the primary visual cortex used direct electrical stimulation at various points along the pathway, or diffuse light flashes. These stimuli have the disadvantage (compared with stimulation by small spots or lines) that they may evoke competing excitatory and inhibitory mechanisms with consequent loss of effectiveness. Nevertheless the responses they do evoke are facilitated by arousal from slow-wave sleep[26–28]. In studies much farther along in the visual pathway, in the inferotemporal region of nitrous oxide-anaesthetized monkeys, neurones have been observed to respond to complex stimuli more vigorously when the EEG is flat than during periods of high-voltage slow waves[29].

Our results suggest that arousal brings about an improvement in signal-to-noise ratio: geniculate cells showed enhanced evoked discharges and most cortical cells showed enhanced responses against a lowered or less chaotic background. (Similar effects were seen in somatosensory cortex of monkeys by Gücer[30].) But more than that, all components of the response were enhanced, suppression of firing as well as increases. Arousal may thus be attended by either depression or elevation of firing rate, depending on stimulus conditions. A cell in an awake animal is more the slave of incoming sensory inputs; in sleep it is more independent, less fettered by sensory constraints, but more at the mercy of events related to the still all-too-mysterious slow waves.

In normal life, even with the eyes open and surveying a complex scene, only a small fraction of cells in the visual cortex is likely to be engaged at any given instant because, for example, a light-dark contour can have only one orientation, can move only in one direction and at one velocity. For a few cells the discharge rate will probably be increased by the stimulus, but the remainder will continue to fire at their spontaneous rate, which will on the whole be depressed by arousal. If the visual cortex is representative of the rest of the brain, then in wakefulness most cells are not active but held in quiet readiness.

The effects of arousal in the small sample of geniculate cells studied were surprisingly strong, in contrast with the effects on cortical cells, most of which showed only a modest increase in response. This suggests that for some cortical cells the enhanced responsiveness of geniculate cells may be transmitted not as more vigorous firing but as more selective firing. Several cells showed a relatively greater enhancement to an optimal stimulus than to a suboptimal one, and consequently an improved selectivity for orientation or movement direction. We never saw a decline in selectivity.

In the cat the transition from the awake state to slow-wave sleep may be gradual, characterized in the EEG by bursts of high-voltage slow waves similar to those seen in continuous full-blown slow-wave sleep, alternating every few seconds with the low-voltage fast activity of arousal. The EEG alone cannot tell us whether this represents repeated transitory changes in arousal level or instead is a steady-state level of low arousal (drowsiness) characterized by a fluctuating EEG. Our single-cell recordings from geniculate and cortex support the former alternative, given the often tight correlation, within a few seconds, between a cell's behaviour and the onset or termination of an episode of slow waves.

Certain brain-stem neurones have long been known to show firing patterns that are closely correlated with levels of arousal[31–39]. Recently attention has focused increasingly on two monoaminergic brain-stem regions, the raphé nuclei and the locus coeruleus. Neurones in both these structures fire more rapidly during periods of increased alertness[40–43]. Like visual neurones, the brainstem neurones show a remarkably close correlation with the EEG, but in the brain stem, changes in firing rate usually precede the EEG changes by several seconds[43,44]. As both pharmacological and physiological studies have suggested that monoaminergic brain-stem neurones are important in controlling arousal, we compared, in a few cats, the effects on visual neurones of stimulating the locus coeruleus to the effects of spontaneous arousal and awakening induced by sensory stimuli. In each case they were the same. The effects of natural arousal from slow-wave sleep on spontaneous and evoked activity could be further exaggerated in an already awake animal by a loud noise or pinch; here again locus coeruleus stimulation had the same effect as a loud noise or a pinch.

The deoxyglucose autoradiographs showed clearly that the depression of activity in slow-wave sleep is different in different layers, being especially profound in 5 and 6. (This is consistent with an observation by Singer *et al.*[19], that cortical cells that could be antidromically activated from tectum or geniculate were particularly likely to be activated by brain-stem stimulation.) Our single-cell recordings had hinted at laminar differences, but it was certainly not true that in sleep suppression of responses occurred in all of the deeper-layer cells, or in none of the upper-layer cells. The autoradiographic differences may not be due entirely to cell-body labelling: only electron microscopy will show whether the radioactivity is in cell bodies, axon terminals, dendrites, or all three.

The fact that some cells in layers 5 and 6 shut down in slow-wave sleep is pertinent to the functioning of the superior colliculus and LGB, for they are the respective major targets of these two layers. One guess might be that the source of the arousal effects observed in the LGB is the brain stem, because the locus coeruleus is known to send strong projections to the geniculate, and the dorsal raphé nucleus probably does[45,46]. Arousal could also act on the geniculate indirectly by reviving some otherwise dormant cells in layer 6 (ref. 19). This could be tested by cortical cooling or injection of local anaesthetics during examination of arousal effects in geniculate cells. It will be especially interesting to compare the responses of cells in the superior colliculus during waking and sleeping, because cooling and ablation of the visual cortex are known to abolish directionally selective responses to moving visual stimuli[47].

Our physiological findings imply either that the responses of many visual neurones in the central nervous system are enhanced on arousal, or that in sleep the responses are dulled. As the function of sleep is unknown, it is, of course, not clear whether a muffling of sensory input during sleep serves to insulate the animal from its environment to permit uninterrupted sleep or whether sensory systems also need to rest and the decreased responsiveness we see reflects whatever recuperative process the brain undergoes during sleep.

David Freeman did the computer programming and electronics, the histology was done by Birthe Storai, Suzanne Fenstemaker and Debra Hamburger, and the photography by Marc Peloquin. We thank David Ferster for interesting criticisms. Simon LeVay suggested the technique of overlaying adjacent autoradiographs. Some of the early work on developing the double label was done with Dr Gary Blasdel. The work was supported by NIH grant R01 EY00605, grants from the Rowland Foundation and the Klingenstein Fund and by NIH grant R01 NS 07848 to E. A. Kravitz.

Received 5 January; accepted 27 March 1981.

1. Hubel, D. H. *J. Physiol., Lond.* **150,** 91–104 (1960).
2. Singer, W. *Physiol. Rev.* **57,** 386–420 (1977).
3. Hubel, D. H. *Am. J. Ophthal.* **46,** 110–122 (1958).
4. Hubel, D. H. *J. Physiol., Lond.* **147,** 226–238 (1959).
5. Evarts, E. V. in *The Nature of Sleep* (eds Wolstenholme, G. E. W. & O'Connor, M.) 171–182 (Churchill, London, 1961).
6. Evarts, E. V., Bental, E., Bihari, B. & Huttenlocher, P. R. *Science* **135,** 726–728 (1962).
7. Coenen, A. U. L. & Vendrik, A. J. H. *Expl Brain Res.* **14,** 227–242 (1972).
8. Hubel, D. H. & Livingstone, M. S. *Neurosci. Abstr. 10th a. Meet.*, Cincinnati, No. 113.3 (1980).
9. Sokoloff, L. *et al. J. Neurochem.* **28,** 897–916 (1979).
10. Livingstone, M. S. & Hubel, D. H. *Neurosci. Abstr. 10th a. Meet.*, Cincinnati, No. 113.4 (1980).
11. Singer, W. & Dräger, U. C. *Brain Res.* **41,** 214–220 (1972).
12. Maffei, L. & Rizzolatti, C. T. *Arch. ital. Biol.* **103,** 609–622 (1965).
13. Sakakura, H. *Jap. J. Physiol.* **18,** 23–42 (1968).
14. Hubel, D. H. & Wiesel, T. N. *J. Physiol., Lond.* **155,** 385–398 (1961).
15. Kuffler, S. *J. Neurophysiol.* **16,** 37–68 (1953).
16. Barlow, H. B. *J. Physiol., Lond.* **119,** 69–88 (1953).
17. Cleland, B. G., Dubin, M. W. & Levick, W. R. *Nature new Biol.* **231,** 191–192 (1971).
18. Meulders, M. & Godfraind, J. M. *Expl Brain Res.* **9,** 201–220 (1969).
19. Singer, W., Tretter, F. & Cynader, M. *Brain Res.* **102,** 71–90 (1976).
20. Hubel, D. H. & Wiesel, T. N. *J. Neurophysiol.* **28,** 229–289 (1965).
21. Hubel, D. H. & Wiesel, T. N. *J. Physiol., Lond.* **160,** 106–154 (1962).
22. Hubel, D. H. & Wiesel, T. N. *J. Physiol., Lond.* **195,** 215–243 (1968).
23. Wurtz, R. H. *J. Neurophysiol.* **32,** 727–742 (1969).
24. Sherrington, C. S. *Man on His Nature*, 183–184 (Doubleday, New York, 1953).
25. Murata, K. & Kameda, K. *Arch. ital. Biol.* **101,** 306–331 (1963).
26. Akimoto, H., Saito, Y. & Nakamura, Y. *Neurophysiologie und Psychophysik des visuellen Systems* (eds Jung, R. & Kornhuber, H.) 363–374 (Springer, Berlin, 1961).
27. Evarts, E. V. *J. Neurophysiol.* **26,** 229–248 (1963).
28. Evarts, E. V. in *Colloque International sur les Aspects Anatomo-fonctionnels de la Physiologie du Sommeil*, 189–212 (Éditions du Centre National de la Recherche Scientifique, Paris, 1965).
29. Gross, C. G., Rocha-Miranda, C. E. & Bender, D. B. *J. Neurophysiol.* **35,** 96–111 (1972).
30. Gücer, G. *Expl Brain Res.* **34,** 287–298 (1979).
31. Mollica, A., Moruzzi, G. & Naquet, R. *EEG J.* **5,** 571–584 (1953).
32. Machne, X., Calma, I. & Magoun, H. W. *J. Neurophysiol.* **18,** 547–558 (1955).
33. Strumwasser, F. *Science* **127,** 469–470 (1958).
34. Schlag, J. *L'activité Spontanée des Cellules du Système Nerveux Central* (Editions Arscia, Bruxelles, 1959).
35. Huttenlocher, P. R. *J. Neurophysiol.* **24,** 451–568 (1961).
36. Beniot, O. *J. Physiol., Paris* **56,** 259–262 (1964).
37. Kasamatsu, T. *Brain Res.* **14,** 506–509 (1969).
38. Kasamatsu, T. *Expl Neurol.* **28,** 450–470 (1970).
39. Hobson, J. A., McCarley, R. W., Pivik, R. T. & Freedman, R. *J. Neurophysiol.* **37,** 497–511 (1974).
40. Chu, N.-S. & Bloom, F. E. *Science* **179,** 908–910 (1973).
41. McGinty, D. J., Harper, R. M. & Fairbanks, M. K. *Serotonin and Behavior* (eds Barchas, S. & Usdinin, E.) 267–279 (Academic, New York, 1973).
42. McGinty, D. J. & Harper, R. M. *Brain Res.* **101,** 569–575 (1976).
43. Foote, S. L., Aston-Jones, G. & Bloom, F. E. *Proc. natn. Acad. Sci. U.S.A.* **77,** 3033–3037 (1980).
44. Trulson, M. E. & Jacobs, B. L. *Brain Res.* **163,** 135–150 (1979).
45. Moore, R. Y. & Bloom, F. E. *A. Rev. Neurosci.* **2,** 113–168 (1979).
46. Moore, R. Y., Halaris, A. E. & Jones, B. E. *J. comp. Neurol.* **180,** 417–438 (1978).
47. Wickelgren, B. G. & Sterling, P. *J. Neurophysiol.* **32,** 16–32 (1969).

LETTERS TO NATURE

Density power spectrum in the local interstellar medium

J. W. Armstrong*, J. M. Cordes† & B. J. Rickett‡

* Jet Propulsion Laboratory, California Institute of Technology,
4800 Oak Grove Drive, Pasadena, California 91109, USA
† Department of Astronomy, Cornell University, Space Science
Building, Ithaca, New York 14853, USA
‡ Department of Electrical Engineering and Computer Sciences,
University of California, San Diego, La Jolla, California 92093, USA

Since the identification of interstellar scintillation, the interstellar medium (ISM) has been known to contain electron density fluctuations having scales $\sim 10^9$ m. The scattering properties of these irregularities have been used in pulsar and continuum source studies and to investigate the small scale turbulence itself[1-11]. (We use 'turbulence' here to imply a disordered, irregular flow.) If the ISM turbulence extends over a broad enough range of fluctuation scales, it could have an important role in galactic cosmic ray transport and confinement[12-14]. Here we present observational evidence that electron density irregularities exist over a very wide range of scale sizes. Radio scattering observations associated with the wavenumber range $\sim 10^{-11}$–10^{-6} m^{-1} are consistent with a density spectrum of the form (wavenumber)$^{-3.7\pm0.6}$. At lower spatial wavenumbers (10^{-16}–10^{-18} m^{-1}) the power spectrum can be estimated by computation of the density fluctuations near the 'outer' scale of the spectrum[15], comparison with density irregularities predicted by theoretical models of the ISM[16], and comparison with observations of the velocity structure function[13,17,18]. When combined with the high-wavenumber (radio) data, these low-wavenumber spectrum estimates are most simply interpreted in terms of a density spectrum behaving as (wavenumber)$^{-3.6\pm0.2}$ over a 12 decade scale size range ($\sim$100 pc to $\sim 10^7$ m). There may be theoretical difficulties[12,14,19] with such a spectrum, however. Alternatively, the low-wavenumber data could be associated with ISM clouds which are physically distinct from the small scale turbulence.

The power spectrum describes the decomposition of a random field's variance into its various Fourier wavenumber components: it partitions the fluctuation 'power' of the field into eddies, with the characteristic size of an eddy being $\sim 2\pi$/wavenumber. The random field of interest here is the interstellar electron density. We assume that the density spectrum P_{3N} (= Fourier transform of the three-dimensional density covariance function) depends only on the modulus of the wavenumber (isotropic statistics) and is independent of position within the ISM (homogeneous statistics). The homogeneity assumption is expected to break down at scale sizes comparable with the overall dimensions of the flow[4]. We have interpreted observations that are sensitive to particular scale sizes in terms of the density fluctuations required on those irregularity scales.

0028-0836/81/250561—04$01.00

Table 1 Estimates of P_{3N}

Method	Wavenumber (m^{-1})	Ref.	Theory equation in ref.	Data ref.	Typical value	C_N^2 ($m^{-6.67}$) (if $\alpha = 11/3$)	$P_{3N}(q)$ (m^{-3}) at $q = 10^{-7}$ m^{-1}
Angular broadening of:							
Crab pulsar	3.8×10^{-8} to 9.2×10^{-7}	4	9	22, 23	Visibility = 0.32 at $\lambda = 4.05$ m	1.7×10^{-4}	7.9×10^{21}
Crab pulsar	$\sim 10^{-8}$ to $\sim 2.4 \times 10^{-7}$	21 4	2.8 19	21, 27, 28	$\theta = 0.1''$ at $\lambda = 4.05$ m	1.2×10^{-4}	5.6×10^{21}
PSR 0834+06		4	19	22, 23	$\theta = 0.06''$ at $\lambda = 2.7$ m	4.4×10^{-4}	2.0×10^{22}
PSR 0950+08		4	19	22, 23	$\theta = 0.05''$ at $\lambda = 2.7$ m	2.2×10^{-3}	1.0×10^{23}
PSR 1133+16	3.8×10^{-8} to 9.2×10^{-7}	4	19	22, 23	$\theta = 0.06''$ at $\lambda = 2.7$ m	1.4×10^{-3}	6.5×10^{22}
PSR 1919+21		4	19	22, 23	$\theta = 0.06''$ at $\lambda = 2.7$ m	7.1×10^{-4}	3.3×10^{22}
IPS survey	$\sim 10^{-8}$ to 2.4×10^{-7}	4	19	25, 26	$\theta = 0.12''$ at $\lambda = 3.5$ m	2.4×10^{-4}	1.1×10^{22}
Visibility phase scintillation	1.2×10^{-8} to 3×10^{-7}	4	9	31	≤ 2 rad^2 at $\lambda = 0.038$ m	≤ 1.5	$\leq 7 \times 10^{25}$ at 10^{-7} m^{-1}
ISS decorrelation bandwidth	$\sim 6 \times 10^{-9}$ to $\sim 6 \times 10^{-8}$	4	41	1, 3, 11	~ 100 kHz at 0.75 m	3×10^{-5} to 3×10^{-4}	$\sim 10^{26}$ at 6×10^{-9} m^{-1}
Pular timing noise	$\sim 2 \times 10^{-12}$ to $\sim 6 \times 10^{-13}$	4 29	28 198	38–41	≤ 1 ms at $\lambda = 2$ m	≤ 4	$\leq 3 \times 10^{43}$ at 2×10^{-12} m^{-1}
Frequency time drift for:							
PSR 1133+16	$\sim 6.3 \times 10^{-12}$ to $\sim 6.3 \times 10^{-10}$	36, 37		9, 11	$\lvert df/dt \rvert \approx 3{,}000$ Hz s^{-1} at 340 MHz	8.2×10^{-4} at $s = 10^{10}$ m	4×10^{30} at 6×10^{-10} m^{-1}
PSR 0329+54		36, 37		11	$\lvert df/dt \rvert \approx 420$ Hz s^{-1} at 408 MHz	9.1×10^{-5} at $s = 10^{10}$ m	5×10^{29} at 6×10^{-10} m^{-1}
PSR 1642−03		36, 37		11	$\lvert df/dt \rvert \geq 50$ Hz s^{-1} at 340 MHz	≤ 4 at $s = 10^{10}$ m	$\leq 2 \times 10^{34}$ at 6×10^{-10} m^{-1}
Velocity structure function	$\sim 3 \times 10^{-18}$ to $\sim 2 \times 10^{-16}$	See text		17, 18	1 km^2 s^{-2} at 1 pc	3×10^{-4} to 3×10^{-8}	6×10^{56} at 10^{-17} m^{-1}
Density outer scale	$\sim 2 \times 10^{-18}$	15, 20 and see text		13, 15 17, 49	$L_0 = 100$ pc $\rho = 0.03$ cm^{-3}	6×10^{-7}	$\sim 5 \times 10^{58}$
ISM clouds	$\sim 10^{-17}$	16		N/A	0.17 cm^{-3}	N/A	$\sim 10^{59}$

This gives a power spectrum level at $q \sim 2\pi/$(scale size) that is consistent with observations. Previous studies[4,8,15] having been confined to fewer types of observational data were limited to estimating the power spectrum in fewer wavenumber ranges. In many wavenumber intervals, we have modelled $P_{3N}(q) = C_N{}^2 q^{-\alpha}$ where C_N^2 measures the 'level' of the turbulence and α is a parameter (usually such that $2 < \alpha < 4$). The P_{3N} estimates from various wavenumber ranges can then be combined to see if an overall form for the spectrum can be synthesized.

Very small scale structure ($\sim 10^6$–10^7 m) can be inferred from observations of the angular broadening of radio sources and very long baseline interferometry (VLBI) phase scintillation. In both observing methods, the phase structure function[4,20] has a central role. In angular broadening observations an interferometer having baseline b measures the visibility $\Gamma(b) = \exp(-D(b)/2)$, where $D(b)$ is the phase structure function: $D(b) = \langle|\phi(\mathbf{x}) - \phi(\mathbf{x}+\mathbf{b})|^2\rangle$. For a spherical wave propagating through a turbulent medium located effectively halfway between the source and the observer, $D(b)$ is related to P_{3N} by

$$D(b) = 8\pi^2 z \lambda^2 r_e^2 \int_0^\infty [1 - J_0(qb/2)] P_{3N}(q) q \, dq \qquad (1)$$

where λ is wavelength, z the source–observer distance, J_0 the Bessel function and r_e the classical electron radius. Thus measurement of $\Gamma(b)$ gives in general only an integral constraint on P_{3N}. However, for a fairly wide class of spectra (for example power laws with $2 < \alpha < 4$) the low wavenumber components of P_{3N} do not contribute to the observation. In these cases, the wavenumber range that mainly determines $\Gamma(b)$ depends on P_{3N}: for $P_{3N} \propto q^{-11/3}$ half of the contribution to $D(b)$ comes from the range 0.1–$2.4\, b^{-1}$ while for $P_{3N} \propto q^{-3}$ this range narrows to 1.6–$5.4\, b^{-1}$. IPS (interplanetary scintillation) observations[21] of the Crab pulsar respond to several effective baselines near 10^6 m and are consistent with $P_{3N} \propto q^{-11/3}$. Hence we associate the spherical wave angular broadening data with the range 0.1–$2.4\, b^{-1}$.

Observations[22,23] at 111 MHz with $b \sim 2.6 \times 10^6$ m give scattering diameters of 0.05–0.07 arc s, yielding estimate of P_{3N} in the range 4×10^{-8}–9×10^{-7} m^{-1} as indicated in Table 1. The IPS method gives information on similar effective baselines[24]. An IPS survey[25] has been interpreted in terms of interstellar scattering[26] and can be incorporated into a power law model framework (see ref. 4 and Table 1). The most detailed IPS observations are of the Crab pulsar[21,27,28]; as with the IPS survey these data give P_{3N} estimates in good agreement with those derived from VLBI data.

Besides angular broadening, the ISM irregularities can cause fluctuations in the apparent position of a radio source[4,29,30] which can be measured with an interferometer. Shapiro *et al.*[31] measured the angular displacement of two radio sources using VLBI, with the observation time short enough that the interferometer was in the same interstellar scintillation (ISS) phase patch[32]. Repeated trials over ~ 3 yr showed a scatter in the angular distance between the sources of $\sim 2 \times 10^{-9}$ rad. This scatter is dominated by other effects (such as tropospheric scintillation), but can be used to construct an upper limit on the the ISM contribution to the interferometric phase fluctuation and hence a limit on P_{3N} in the range 0.05–$1.2\, b^{-1}$ (plane wave). Interpreting 2×10^{-9} rad as an r.m.s., the phase structure function on the $\sim 4 \times 10^6$ m baseline is ~ 2 rad^2. Using the theory[4] with a propagation distance[33,34] ~ 760 pc gives $C_N^2 \leq 1.5$ $m^{-6.67}$ (if $\alpha = 11/3$) and a corresponding P_{3N} limit that is consistent with the above angular broadening measurements.

In principle, pulse broadening[2,6] can also be used to estimate P_{3N} at these large wavenumbers. However, several kpc long propagation distance is required to obtain a measurable effect. Because the ISM turbulence is inhomogeneous on these scales we do not try to incorporate pulse broadening data here.

For scales $\sim 10^9$ m the simplest way to sense ISM eddies is through intensity fluctuations in time[5,35] on frequency–time[1,3,9,11]. In this method, ISM eddies larger than about the Fresnel zone size do not contribute to the observation[4,21,24,29]; signal-to-noise considerations effectively limit the extent to which very small scales can be sensed. By measuring the autocorrelation of the intensity fluctuations in radio frequency, one can get the level and shape of P_{3N} near $q = 10^{-8}$ m^{-1}. These observations[11] favour power law models with α slightly less than 4. Table 1 shows the range of P_{3N} that are consistent with the measured ISS decorrelation bandwidths of relatively nearby pulsars.

Scale sizes between ~10^{10} and 10^{12} m can be detected either through time-of-arrival, fluctuations of pulsar pulses[4,29] or, indirectly, through their effect on the frequency–time structure of the intensity ISS caused by smaller scales[36,37]. The time of arrival fluctuations are caused by time-varying group delay as the blobs move through the pulsar–Earth line. Careful timing observations[38–40] have yielded only a ~1 ms upper limit to ISS time of arrival fluctuations at $\lambda \approx 2$ m and on time scales ~1–3 yr (that is typical spatial scales 3×10^{12}–10^{13} m). For a propagation distance of 1,000 pc we obtain P_{3N} limits as indicated in Table 1.

Alternatively, one can estimate P_{3N} on these wavenumbers through the influence of the large scale blobs on the small scale intensity ISS. The theory is best formulated for two discrete scales[36] and we use it to obtain consistent values of P_{3N} near $q = 10^{-11}\ \mathrm{m}^{-1}$. The physical idea is that, owing to transverse gradients in the phase, the large blobs cause refractive steering of the intensity diffraction pattern of the small (~10^9 m) blobs. The frequency dependence of the plasma refractive index means that there will be a frequency dependent spatial offset of the intensity pattern. If the pattern moves across the Earth with a constant velocity (for the ~1 h of a typical observation) these spatial offsets at different frequencies are manifested as temporal offsets in a frequency–time spectrum (see Fig. 2 of ref. 11). Theory allows computation of the angle of refraction, θ_R, of the large blobs from the slope (df/dt) of the drifting features in frequency–time. Because we are dealing with a discrete scale model, we associate θ_R with the transverse phase gradient[36] ($\theta_R = (\lambda/2\pi)\cdot|\text{grad}\ \phi|$) and then to square-law phase structure function[29] ($|\text{grad}\ \phi| \approx D^{1/2}(b)/b$, where b is a typical refractive blob scale size. The structure function's relationship to P_{3N} is shown for three pulsars in Table 1. Because our theory is really only for two discrete scales the existence of drifting features does not necessarily imply a continuum of scales between ~10^{10} and 10^{12} m.

Some observational aspects of these high-wavenumber P_{3N} estimates should be noted. First, differential VLBI is within two orders of magnitude of being limited by ISS in relative source position measurements. Second, searches for gravitational radiation having periods 1–10 yr using pulsar timing data[41] are apparently limited by pulsar timing noise rather than ISS. (In the Solar System, searches for 30–3,000-s period gravity waves[42,43] are limited by interplanetary phase scintillation[44].) Third, in addition to dispersion measure time of arrival fluctuations, the two-scale theory[36] admits the possibility of time of arrival variations $\tau_R \simeq z\theta_R^2/2c$ (z = typical distance to a scattering element) associated with the changing 'ray' path length to the pulsar. If $\theta_R \sim \theta_S$ (the diffraction scattering angle) then τ_R can be comparable with $z\theta_S^2/2c$ (~a few tens of milliseconds for a highly dispersed pulsar) and hence observable.

The radio data alone are consistent with a pure power law spectrum, $P_{3N} \propto q^{-3.7\pm0.6}$ for $10^{-11} \lesssim q \lesssim 10^{-6}\ \mathrm{m}^{-1}$. This includes both the Kolmogorov spectrum ($\alpha = 11/3$) and the $\alpha = 7/2$ spectrum predicted by dimensionality relations for magnetoacoustic turbulence[45,46]. Both spectra imply a cascade of energy between wavenumbers as in a turbulence process. However, the radio data alone cannot exclude the case $\alpha = 4$ for which the physical explanation is very different. A random assembly of density discontinuities would contribute a q^{-4} spectrum at high wavenumbers; such a spectrum does not imply any energy cascade. However, where sufficiently accurate ISS radio data have been recorded[11], the spectra near a wavenumber $q \sim 10^{-9}\ \mathrm{m}^{-1}$ are consistent with a power slightly less than 4. Further, if the estimates near $q \sim 10^{-18}\ \mathrm{m}^{-1}$ discussed below are correct, the data are consistent with a $(\text{wavenumber})^{-3.6\pm0.2}$ power law over 12 orders of magnitude in wavenumber and exclude the $\alpha = 4$ case.

Low wavenumber estimates can be obtained by computing the density fluctuations near the 'outer' scale of the turbulence, by interpreting the velocity structure function data, and through theoretical predictions for the large-scale density fluctuations in the ISM. The outer scale argument is quite general[15,20] and as the density structure function $D_N(\mathbf{r})$ is the mean square of the difference in density at two points, $D_N(L_0)$ is expected to be comparable to the square of the difference in average density at points separated by the outer scale size L_0, thus $D_N(L_0) \sim (L_0\ \text{grad}\ \bar{\rho})^2$. If $D_N(\mathbf{r}) = (C_N^2/0.033)r^{2/3}$ then $C_N^2 \sim 0.033\ L_0^{4/3}\ (\text{grad}\ \bar{\rho})^2$, where the factor 0.033 is a consequence of our spectrum definition (see refs 4 and 20). Taking $\bar{\rho} = 0.03\ \mathrm{cm}^{-3}$, $L_0 =$ 100 pc (refs 15, 17), and the gradient scale for $\bar{\rho}$ to be 500 pc (ref. 34) gives $C_N^2 \sim 10^{-6.2}\ \mathrm{m}^{-6.67}$. The associated value of $P_{3N}(2\pi/L_0)$ is plotted in Fig. 1 with an error of ± two orders of magnitude.

Velocity structure function data give further evidence for a turbulence-type power law spectrum at low wavenumbers. The velocity structure function $D_v(\mathbf{r}) = \langle|v(\mathbf{s}) - v(\mathbf{s}+\mathbf{r})|^2\rangle$ is measured[17,18] from the motion of both stars and neutral gas to be $D_v \propto r^{2/3}$ (which corresponds to the Kolmogorov spectrum) for scales 1–100 pc (Fig. 16 of ref. 17 and Fig. 1 of ref. 18 are particularly striking in this regard). To relate the level of the velocity fluctuations to the level of density fluctuations (hence P_{3N}) we have to consider a model of the ISM. There is now strong evidence for a coronal phase (10^6 K, diffuse, ionized) in the ISM[47]. Our interpretation is based on a specific model[16] with three main components: cool (~10^2 K) dense neutral gas clouds surrounded by a warm (~10^4 K) gas partially ionized by stellar UV, embedded in the hot coronal phase driven by supernova expansions. The coronal phase occupies a large fraction of the space but only contributes a small fraction of the ionized column density on typical pulsar distances. Thus at low wavenumbers $P_{3N}(q)$ probably reflects the warm phase which is thought to be closely associated with the cool neutral phase. The neutral gas motions, $D_v \propto r^{2/3}$, are also typical of the warm ionized phase and imply a Kolmogorov spectrum shape near $q \sim 3\times10^{-18}$–$2\times10^{-16}\ \mathrm{m}^{-1}$. It is not clear, however, how to relate the magni-

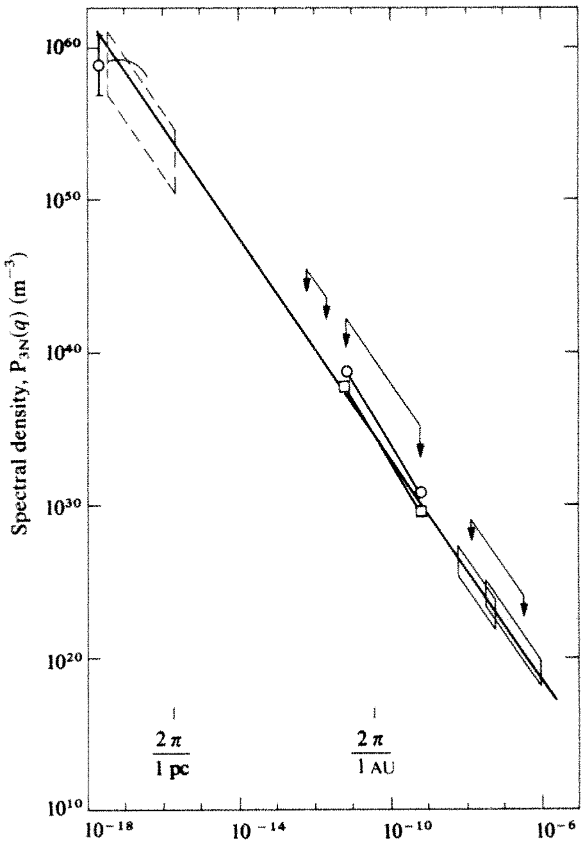


Fig. 1 Density power spectrum in the local interstellar medium. Spectral estimates are obtained from Table 1. Starting at the lowest wavenumber, estimates or limits come from: density outer scale, ISM clouds, velocity structure function, pulsar timing noise, frequency–time drift in dynamic spectra, intensity scintillation, visibility phase scintillation, and angular broadening. The line through the data has a logarithmic slope of −11/3 (Kolmogorov spectrum); slope of −3.5 is also consistent with the data.

tude of the velocity fluctuations to the level of P_{3N} for the density fluctuations in the warm partially ionized plasma. The estimate included here is based on analogy with the solar wind plasma, but it may not be physically justified for the ISM. For scales which are small compared with the overall dimensions of the solar wind flow, the fractional density fluctuations are proportional[48] to the velocity fluctuations divided by the Alfvén speed: $|\delta\rho/\bar{\rho}| = a|\delta v/v_A|$ where $a = 10^{0\pm1}$. We have converted these data to P_{3N} as shown in Table 1, using $\bar{\rho} = 0.03\ \mathrm{cm}^{-3}$ and $v_A = 10\ \mathrm{km\ s}^{-1}$. Because of the uncertainty in relating the levels of the velocity and density fluctuations, the P_{3N} estimates are shown in Fig. 1 as a dashed box. The vertical extent of the box is the ± two decade uncertainty associated with a.

Another low-wavenumber estimate is based on the density inhomogeneities predicted by the above ISM model. The model predicts that a significant fraction of space is filled with warm partly ionized cloud shells having typical electron densities $\sim 0.17\ \mathrm{cm}^{-3}$ and typical radii ~2.1 pc. If these are approximated as spheres having gaussian radial density profiles, we obtain the spectrum indicated in Fig. 1, plotted over a decade of wavenumber near $q = 10^{-17}\ \mathrm{m}^{-1}$. This estimate agrees well with both the outer scale point and the spectrum inferred from the velocity data.

The striking feature of Fig. 1 is that the data are most simply represented as proportional to $q^{-\alpha}$ with $3.4 \leqslant \alpha \leqslant 3.8$ over 12 orders of magnitude in q. These conclusions are in general agreement with those of other workers[13,15,37,49], although the present data set is more extensive. However, there may be physical difficulties with an inertial subrange-type spectrum over this very wide scale size range. The excitation and damping of various wave modes in the ISM have been discussed elsewhere[12,14,19,48,51]. It is expected that the coronal phase could[12] support a power law-type spectrum between $10^{-17}\ \mathrm{m}^{-1} \leqslant q \leqslant 10^{-9}\ \mathrm{m}^{-1}$, with higher wavenumber modes damped. The warm partly ionized component, on the other hand, is thought[14] only capable of supporting fluctuations with $q \leqslant 10^{-13}\ \mathrm{m}^{-1}$. Finally, the cool nearly neutral component is predicted[19] to contain only high-wavenumber ($q \gtrsim 10^{-9}$) irregularities. If these theoretical difficulties are real, different wavenumber ranges in Fig. 1 must be associated with different ISM components. High-wavenumber pulsar scintillations and the pulsar dispersion measure might then be linked only statistically through the distance. An alternative picture[50] of the ISM associates dispersing electrons with medium density Strömgren spheres of a space density much lower than that of the warm clouds in the model[16] discussed above. Further investigations into the scale heights of the dispersing and (high-wavenumber) scattering electrons will reveal whether they exist in the same phase. If different ISM components contribute to different wavenumber ranges of P_{3N}, the very low wavenumber Fourier components might be associated with warm, partially ionized clouds that truly have a characteristic scale. The large scale fluctuations could then be physically distinct from the high wavenumber turbulence. This view requires an approximately power law spectrum over a much smaller range of wavenumbers and thus partially obviates the theoretical objections noted above. The lack of observations in the 10^{-16}–$10^{-12}\ \mathrm{m}^{-1}$ range does not rule out this interpretation; however, it seem at odds with the continuum of scales suggested by the shape of the velocity structure function.

Our estimate of a global average wavenumber spectrum for the interstellar electron density is most simply represented by a power law $\propto q^{-\alpha}$ with $3.4 \leqslant \alpha \leqslant 3.8$ over $10^{-18} \leqslant q \leqslant 10^{-6}\ \mathrm{m}^{-1}$, suggesting energy exchange over 12 orders of magnitude in scale. There are theoretical arguments against such a view based on current studies of wave modes in the warm and hot phases of the ISM. Clearly future work must reconcile the results of Fig. 1 with models of the ISM and also examine the possibility of energy exchange between irregularity scales over a very wide range.

We thank J. R. Jokipii for interesting discussions and for bringing ref. 17 to our attention. This work was supported in part by the Caltech President's Fund, NASA contract NAS 7-100, NSF grant ATM 8012801, and the National Astronomy and Ionosphere Center which is operated by Cornell University under contract with the NSF.

Received 4 March; accepted 16 April 1981.

1. Rickett, B. J. *Mon. Not. R. astr. Soc.* **150,** 67–91 (1970).
2. Sutton, J. M. *Mon. Not. R. astr. Soc.* **155,** 51–64 (1971).
3. Roberts, J. A. Ables, J. G. Preprint (1980).
4. Rickett, B. J. *A. Rev. astr. Astrophys.* **15,** 479–504 (1977).
5. Backer, D. C. *Astrophys. J.* **190,** 667–671 (1974).
6. Komesaroff, M. M., Hamilton, P. A. & Ables, J. G. *Aust. J. Phys.* **25,** 759–777 (1972).
7. Rickett, B. J. & Lang, K. R. *Astrophys. J.* **185,** 945–950 (1973).
8. Mutel, R. L. *et al. Astrophys. J.* **193,** 279–282 (1974).
9. Ewing, M. S., Batchelor, R. A., Friefeld, R. D., Price, R. M. & Staelin, D. H. *Astrophys. J.* **162,** L169–L172 (1970).
10. Slee, O. B. *et al. Mon. Not. R. astr. Soc.* **167,** 31–47 (1974).
11. Armstrong, J. W. & Rickett, B. J. *Mon. Not. R. astr. Soc.* **194,** 623–638 (1981).
12. Cesarsky, C. J. *A. Rev. Astr. Astrophys.* **18,** 289–319 (1980).
13. Jokipii, J. R. *Interaction of Charged Particles with Random Electromagnetic Fields in Astrophysics* (Summer school on Plasma Instabilities and Turbulence, Aix-en-Provence, 1974).
14. McIvor, I. *Mon. Not. R. astr. Soc.* **179,** (1977); **178,** 85–99 (1977).
15. Lee, L. C. & Jokipii, J. R. *Astrophys. J.* **206,** 735–743 (1976).
16. McKee, C. F. & Ostriker, J. P. *Astrophys. J.* **218,** 148–169 (1977).
17. Kaplan, S. A. *Interstellar Gas Dynamics* (Pergamon, Oxford, 1966).
18. Larson, R. B. *Mon. Not. R. astr. Soc.* **186,** 479–490 (1979).
19. Scheuer, P. A. G. & Tsytovich, V. N. *Astrophys. Lett.* **7,** 125–127 (1970).
20. Ishimaru, I. *Wave Propagation and Scattering in Random Media* Vol. 2, 538 (Academic, New York, 1978).
21. Kaufman, J. J. thesis, Univ. California (1976).
22. Vandenberg, N. R. thesis, Univ. Maryland (1974).
23. Resch, G. M. thesis, Florida State Univ. (1974).
24. Cohen, M. H., Gundermann, E. J., Hardebeck, H. E. & Sharp, L. E. *Astrophys. J.* **147,** 449 (1967).
25. Readhead, A. C. S. & Hewish, A. *Mem. R. astr. Soc.* **78,** 1 (1974).
26. Duffett-Smith, P. J. & Readhead, A. C. S. *Mon. Not. R. astr. Soc.* **174,** 7–17 (1976).
27. Coles, W. A. & Kaufman, J. J. *Mon. Not. R. astr. Soc.* **181,** 57P–59P (1977).
28. Bell, S. J. & Hewish, A. *Nature* **213,** 1214 (1967).
29. Lovelace, R. V. E. thesis, Cornell Univ. (1970).
30. Cronyn, W. M. *Astrophys. J.* **174,** 181–200 (1972).
31. Shapiro, I. I. *et al. Astr. J.* **84,** 1459–1469 (1979).
32. Cohen, M. H. & Cronyn, W. M. *Astrophys. J.* **192,** 193–197 (1974).
33. Taylor, J. H. & Manchester, R. N. *Astrophys. J.* **215,** 885–896 (1977).
34. Readhead, A. C. S. & Duffett-Smith, P. J. *Astr. Astrophys.* **42,** 151–153 (1975).
35. Downs, G. S. & Reichley, P. E. *Astrophys. J. Lett.* **163,** L11–L16 (1971).
36. Shishov, V. I. *Soviet Astr.* **17,** 598–602 (1974).
37. Hewish, A. *Mon. Not. R. astr. Soc.* **192,** 799 (1980).
38. Helfand, D. J., Taylor, J. H., Backus, P. R. & Cordes, J. M. *Astrophys. J.* **237,** 206–215 (1980).
39. Cordes, J. M. & Helfand, D. J. *Astrophys. J.* **239,** 640–650 (1980).
40. Gullahorn, G. E. & Rankin, J. M. *Astr. J.* 1219 (1978).
41. Detweiler, S. *Astrophys. J.* **234,** 1100–1104 (1979).
42. Estabrook, F. B. & Wahlquist, H. D. *Gen. rel. Gravit.* **6,** 439 (1975).
43. Wahlquist, H. D., Anderson, J. D., Estabrook, F. B. & Thorne, K. S. *Atti dei Convengui Lincei,* **34,** 335 (1977).
44. Armstrong, J. W., Woo. R. & Estabrook, F. B. *Astrophys. J.* **230,** 570–574 (1979).
45. Kraichnan, R. H. *Phys. Fluids* **8,** 1385 (1965).
46. Chashei, I. V. & Shishov, V. I. *Geomagn. Aeron.* **17,** 662–667 (1977).
47. McCray, R. & Snow, T. P. *A. Rev. Astr. Astrophys.* **17,** 213–240 (1979).
48. Neugebauer, M., Wu, C. S. & Huba, J. D. *J. geophys. Res.* **83,** 1027–1034 (1978).
49. Jokipii, J. R. & Lerche, I. *Astrophys. J.* **157,** 1137 (1969).
50. Salpeter, E. E. in *The Large Scale Characteristics of the Galaxy* (ed. Burton, W. B.) 245–252 (Reidel, Dordrecht 1979).
51. Hall, A. N. *Mon. Not. R. astr. Soc.* **190,** 353–369 (1980).

Atmospheric nitrous oxide produced by solar protons and relativistic electrons

Sheo S. Prasad

Jet Propulsion Laboratory, California Institute of Technology,
4800 Oak Grove Drive, Pasadena, California 91109, USA

E. C. Zipf

University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA

Nitric oxide (NO), which contributes to the destruction of stratospheric ozone, may be formed directly in the upper atmosphere by solar protons[1,2] and by the precipitation of relativistic electrons from the Earth's radiation belts[3]. We now describe an alternative means by which solar proton (SP) events and relativistic electron precipitation (REP) events may lead to the production of stratospheric NO—the production of nitrous

0028–0836/81/250564—03$01.00

oxide (N_2O) in the mesosphere, its downward migration and conversion in the stratosphere to NO by the reaction

$$N_2O + O(^1D) \rightarrow 2\ NO \qquad (1)$$

This process could amplify the direct NO production by >10%, which is significant. Mesospheric nitrous oxide mixing ratios increase to values as high as 6×10^{-7}. due to REP- and SP-related production. Lateral transport will reduce these high values. But even so, mesospheric mixing ratios of N_2O in the high latitudes would approach 10^{-7}, which is considerably greater than those expected on the basis of theories which neglect REP- and SP-related production of this species.

Production of N_2O by precipitating energetic particles is a consequence of the following two reactions[4–8].

$$N_2 + e^- \rightarrow N_2(A^3\Sigma) + e^- \qquad (2)$$

$$N_2(A^3\Sigma) + O_2 \rightarrow N_2O + O \qquad (3)$$

where e^- represents the primary and secondary electrons. Although the amount of $N_2(A^3\Sigma)$ per ion pair is uncertain, a value of 0.52 $N_2(A^3\Sigma)$ per ion pair in air seems most probable on the basis of most recent electron impact cross-sectional data of Cartwright *et al.*[9] as interpreted by Gilmore[10]. Zipf[5] has experimentally determined a 60% probability for N_2O production in the quenching of $N_2(A^3\Sigma)$ by O_2 which is in good agreement with earlier experimental results of Malcombe-Lawis[4]. Furthermore, chemical reaction (2) and physical quenching by O_2 seem to be the sole fate of $N_2(A^3\Sigma)$ below about 90 km (refs 7, 8). Thus, a value of 0.31 N_2O per ion pair has been adopted in this study.

Production of N_2O by energetic particles can now be obtained directly from the ion production rates due to these particles. Instead of considering an isolated SP or REP event, we are considering the collective effects of all particle precipitation events, both large and small. This has been done in a simplified manner by using yearly averages of ion production rates. Figure 1 shows the results for energetic solar protons and relativistic electrons. Production rates for solar protons are based on the average ion production rate calculated by C. H. Jackman (personal communication) for the year 1978 on the basis of SP observations by experiment on board IMP 8. For relativistic electrons we used the estimates of ion production rates given by Thorne[11]. We chose to use the ion production rate corresponding to REP occurrence probability of 10% rather than 1%, because recent experiments using heavy-duty magnetic spectrometers with a large geometric factor flown on S3-2 and bremsstrahlung X-ray imaging experiments flown on 1972-076B satellites suggest that the occurrence probability, as well as the fractional area of the globe affected by REP events, are significantly larger than their current estimates[12]. Thorne's[11] ion production rates are for the subauroral zone ($4 \leq L \leq 10$). Possibly similar production rates may be valid for some higher latitudes also. Nitrous oxide production rates used in our calculations are the sums of the two rates shown in Fig. 1.

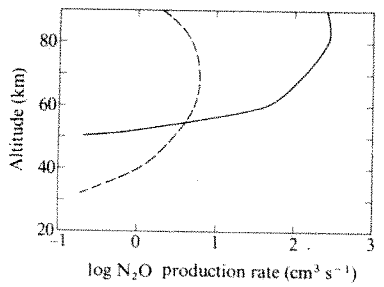


Fig. 1 N_2O production rates: ---, average value for the year 1978 due to solar protons; ——, due to relativistic electrons assuming REP events occur 10% of the time.

Atmospheric number densities and mixing ratios of N_2O corresponding to these sources were calculated using one-dimensional photochemical–dynamical modelling techniques[13]. We assumed that N_2O molecules are lost by photodissociation and by reaction with $O(^1D)$, and by ionic reactions in the ionosphere. Concentrations of $O(^1D)$ were calculated on the basis of local chemical equilibrium between its production from O_3 photolysis and loss by quenching collisions with N_2 and O_2. The empirical method of Shimazaki *et al*[14]. was used to determine the absorption of solar UV in the Schuman–Runge bands. We assumed an N_2O mixing ratio of 3×10^{-7} at the lower boundary and a zero flux at the upper boundary at 100 km. Our model was verified by making it reproduce currently believed theoretical distributions[15] of N_2O in the high latitudes (~60°) when the flux from the Earth's surface is assumed to be the only source. Results from calculations done for a subauroral latitude of 65° and equinoxial conditions are illustrated in Figs 2 and 3.

From the data presented in Fig. 2, it is clear that nitrous oxide production by REP and SP events could lead to substantial enhancements in high-latitude stratospheric and mesospheric N_2O mixing ratios. These mixing ratios reach values as high as $(3–5)\times10^{-7}$ in the 70–80 km region in the present one-dimensional calculation. Introduction of meridional transport will reduce these numbers. But even so, mesospheric N_2O mixing ratios in the high latitudes would remain substantially greater than those expected on the basis of theories which neglect REP- and SP-related production. This was verified by an approximate calculation in which N_2O loss rate coefficients were systematically increased to simulate roughly the effects of meridional flow.

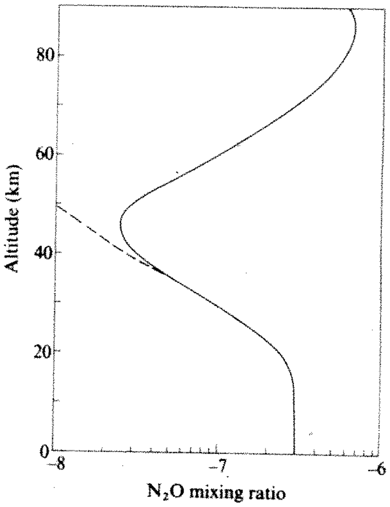


Fig. 2 Theoretical N_2O mixing ratios in the high latitudes neglecting horizontal transport: ——, with the N_2O production due to energetic particles; ---, without this source.

Indirect column production rates of NO for the various altitudes (h) through reaction (1), the integral $2k_1 \int_h^{100\ \mathrm{km}} n(N_2O)n(O^1D)\,dh$, are shown in Fig. 3, where curves *a* and *b* represent the column production rates for the two cases of with and without the particle precipitation related N_2O source. The indirect production of NO through reaction (1) should be distinguished from the direct production of NO discussed by Crutzen *et al.*[1] and Thorne[3]. The column production rates for this direct case are shown by curve *c* in Fig. 3 for comparison with the indirect case. In either case, the production rate of NO in any given region of the atmosphere bounded by altitudes h_1 and h_2 ($h_2 > h_1$) can be obtained by simply subtracting the column production rate at h_2 from that at h_1. In this manner, the data presented by curves *a* and *b* of Fig. 3 show that the indirect

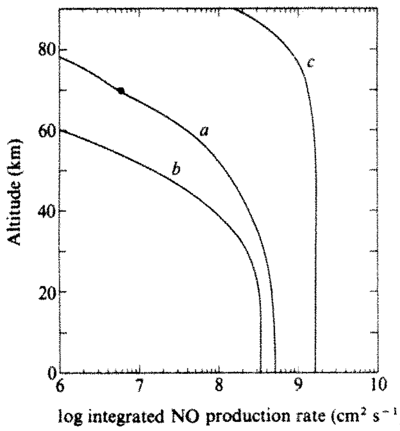


Fig. 3 Column production rate of NO. *a*, Indirect production using N_2O profile obtained with the energetic particle related source of nitrous oxide; *b*, the same without this source; *c*, for direct production of NO by energetic particles.

production of NO in the stratosphere and mesosphere due to precipitating particles was $1.6 \times 10^8 \text{ cm}^{-2} \text{ s}^{-1}$ in our model. Direct production of NO in the same column was $1.7 \times 10^9 \text{ cm}^{-2} \text{ s}^{-1}$ (see curve *c*).Thus, the N_2O mechanism of indirect NO production amplified the direct NO production by ~10%. Furthermore, most of the indirect production of NO is below 60 km, that is, $1.2 \times 10^8 \text{ cm}^{-2} \text{ s}^{-1}$ out of the total $1.6 \times 10^8 \text{ cm}^{-2} \text{ s}^{-1}$ or 75%. In contrast, 75% of the direct production is above 72 km in the present model, where photodissociation rates are higher compared with the rates at and below 60 km (ref. 16). Consequently, only a part of the directly produced NO can be transported below 60 km. Thus, the indirect production of NO through the N_2O mechanism could be more significant than what is implied by the 10% amplification mentioned earlier.

Finally, note that present methods of remotely measuring N_2O from the satellite platform, such as used by the Stratospheric and Mesospheric Sounder[17] on board Nimbus-7, are not considered effective above about 55 km, where measurements are needed most to validate the present theory. Alternative means of measuring high altitude N_2O, for example, from rocket-borne sampling devices, should be considered.

We thank Dr C. H. Jackman for his ion production rate calculation. The contributions of S.S.P. represent one phase of research carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract NAS7-100, sponsored by NASA.

Received 3 December 1980; accepted 10 April 1981.

1. Crutzen, P. J., Isaksen, I. S. A. & Reid, G. C. *Science* **189,** 457 (1975).
2. Reid, G. C., Isaksen, I. S. A., Holzer, T. E. & Crutzen, P. J. *Nature* **259,** 177 (1976).
3. Thorne, R. M. *Science* **195,** 287 (1977).
4. Malcombe-Lewis, D. J. *Nature* **247,** 540 (1974).
5. Zipf, E. C. *Nature* **287,** 523 (1980).
6. Prasad, S. S. in *Solar-Terrestrial Influence on Weather and Climate* (eds McCormac, B. M. & Seliga, T. A.) 299–304 (Reidel, Boston, 1979).
7. Zipf, E. C. & Prasad, S. S. *Nature* **287,** 525 (1980).
8. Swider, W. *Geophys. Res. Lett.* **6,** 335 (1976).
9. Cartwright, D. C., Chutjian, A. & Trajmar, S. *Phys. Rev.*, **A16,** 1041 (1977).
10. Gilmore, F. *Pap. at Defense Nuclear Agency-Air Force Geophysical Laboratory Nuclear Weapons Effects Chemistry Conf.* Bedford (1980).
11. Thorne, R. M. *Pure appl. Geophys.* **118,** 128 (1980).
12. Imholf, W. L., Nightingale, R. W., Reagan, J. B. & Nakano, G. H. *J. atmos. terr. Phys.* **42,** 443 (1980).
13. Turco, R. P. & Whitten, R. C. *NASA tech Pap.* 1002 (NASA, Washington DC, 1977).
14. Shimazaki, T., Ogawa, T. & Farrel, B. C. *NASA tech. Not.* TND-8399 (NASA, Washington DC, 1977).
15. Logan, J. A., Prather, M., Wofsy, S. & McElroy, M. B. *Phil. Trans. R. Soc.* **A290,** 187 (1978).
16. Nicolet, M. & Cieslik, S. *Planet. Space Sci.* **28,** 105 (1980).
17. Drummond, J. R. *et al. The Nimbus 7 User's Guide* (ed. Madrid, C. R.) (Goddard Space Flight Center, Greenbelt).

Field and structure behaviour of electrolytes

James W. Cobble & Richard C. Murray Jr

Department of Chemistry, San Diego State University, San Diego, California 92182, USA

Utpal Sen

Central Electrochemical Research Institute, Karaikudi 623006, Tamil Nadu, India

Measurements of the thermodynamic properties of electrolytes in hot water have a widespread practical and theoretical interest. The chemistry of many solutes changes rapidly as a solution is heated but the reasons for such changes have not been understood in terms of electrolyte theory. The temperature behaviour of the equilibrium thermodynamic properties is best seen by examining the ionic heat capacities of representative electrolytes, and we report here measurements just completed of $\bar{C}^\circ_{p2}$ of NaCl(aq) and $BaCl_2$ (aq) up to 300 °C. The values are so negative that we now believe that the properties of electrolytes at higher temperatures are determined primarily by the electrostatic field of the ions and the bulk dielectric constant of water, and not by structural arrangements of ions and water molecules which characterize solutions at lower temperatures. These new phenomena are summarized in Fig. 1, which not only illustrates the large values at 300 °C for $\bar{C}^\circ_{p2}$ for NaCl (aq) and $BaCl_2$ (aq) of −339 and −1,011 cal mol^{-1} deg^{-1}, but also indicates a continued rapid change above 300 °C. The observations also suggest a need for reexamination of the theory of aqueous electrolytes at 25 °C, a temperature in the transition region between the two types of behaviour.

The standard state heat capacities for the electrolytes were obtained in a pressurized titanium–palladium calorimeter using the integral heat method[1] of dissolving salts to give very dilute solutions (down to 5×10^{-5} molal) and extrapolating the resulting heats of solution to infinite dilution. These heats at 10 temperatures between 50 and 300 °C were corrected to a pressure of 1 atm for calculation, using the molal volume data of Ellis and McFadden[2] up to 200 °C and extrapolated values up to 300 °C. The corrected data were fitted to a five-term power

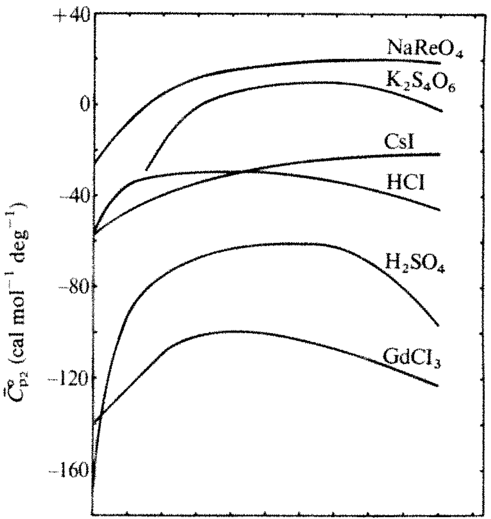


Fig. 1 Standard-state partial molal heat capacity of *a*, NaCl, and *b*, $BaCl_2$, 0–300 °C (solid heavy line) at P_{sat}. The Born electrostatic heat capacity function is indicated by the band.

0028–0836/81/250566—03$01.00

Table 1 Comparison of 'field' and 'structure' electrolyte behaviour

Structure	Field
1. $(\mathrm{d}\bar{C}^{\circ}_{p2}/\mathrm{d}T) > 0$	$(\mathrm{d}\bar{C}^{\circ}_{p2}/\mathrm{d}T) < 0$
2. Primary and secondary hydration of ions	Continuum of hydration around ions
3. Properties affected by structure of solvent	Properties dependent only on field of ion and *D* of solvent
4. Unequal hydration of cations and anions	
5. Structure-making and structure-breaking ions	

series and differentiated to obtain ΔC_p; $\bar{C}^{\circ}_{p2}$ was then obtained by addition of the known heat capacities of the solids[3] over the same temperature range. The 1-atm electrolyte heat capacities were then corrected back to the solvent pressure at each temperature (Fig. 1). The pressure corrections are negligible below 200 °C and were <4% even at the highest temperatures. The heat capacities below 100 °C for $BaCl_2$ and 200 °C for NaCl are from previously published data[1,4]. The use of very dilute solutions is important as ionic strength effects are much larger at elevated temperatures. For example, the heat capacity of dilution of a 0.006 molal $BaCl_2$ solution to infinite dilution from the present data is roughly −250 cal mol^{-1} deg^{-1} at 300 °C. The data obtained by flow calorimetry, which is limited to higher concentrations ($>0.1m$), also illustrates this point. Thus, Wood and Smith-Magowan[5] indicate a heat capacity of dilution of NaCl from $m = 0.2433$ to $m = 0.0996$ of approximately −100 cal mol^{-1} deg^{-1}, at 175 atm. The heat capacity function, although difficult to obtain, is more temperature sensitive than other molal thermodynamic functions which are obtained by integration of $\bar{C}^{\circ}_{p2}$, a process which conceals many of the details of ion–solvent interactions.

Most investigators believe that the properties of water[6–12] and aqueous solutions[6,12–16] near room temperature depend heavily on the hydrogen-bonded structure of the solvent. Competition between an ion and the normal solvent structure for solvent molecules led Frank and Wen[13] to postulate three regions of solvation surrounding a central ion: the nearest-neighbour hydration sphere consisting of water molecules which are mostly, if not completely[16], oriented towards the centre of the ionic charge; a second, transitional zone of non-structured water molecules; and finally, the outer zone comprised of the normal, structured 'bulk' water. However, on heating to higher temperatures, the hydrogen-bonded structure of bulk water begins to break down[12–17] and the dielectric constant, *D*, falls from 87.85 at 0 °C to 20.05 at 300 °C, although the (*DT*) product is still large enough to sustain ionic behaviour[8]. The effects of the field of an ion, which, at 25 °C, are limited to small distances as a result of the high solvent dielectric constant and the structure built around the ion, should extend much further out from the ion and into the bulk solvent at 300 °C. The very negative values of $\bar{C}^{\circ}_{p2}$ obviously reflect the loss of degrees of freedom of increasing numbers of water molecules at higher temperatures as the following estimate illustrates. If $\bar{C}^{\circ}_{p}(\mathrm{Na}^{+}) \approx \bar{C}^{\circ}_{p}(\mathrm{Cl}^{-}) \approx -170$, then $\bar{C}^{\circ}_{p}(\mathrm{Ba}^{+2}) \approx -670$ cal mol^{-1} deg^{-1}. Estimating that each water molecule involved in hydration loses, on the average, about one-quarter of its normal heat capacity (in tightly bound crystalline salts the loss is about one-half), we can easily calculate that Ba^{2+} influences 670 ÷ 6.18, or 109, water molecules at 300 °C; at 25 °C only 12 molecules are influenced. At 300 °C, the bulk density of water indicates an effective radius of 2.1 Å for a water molecule. Assuming a primary coordination number of six, the distance of influence extending out from the centre of Ba^{2+} corresponding to 109 water molecules is 13 Å (and similarly, 9 Å for Na^{+}).

The low-temperature parts of the heat capacity curves in Fig. 1 indicate an opposite behaviour, where $\bar{C}^{\circ}_{p2}$ values become more negative with decreasing temperature, undoubtedly due to increasing solvent structural effects around the ion.

We propose the terms 'structure' (dominated) and 'field' (dominated) electrolytes to distinguish between these extremes in behaviour; in the former the structural effects of the solvent and its interaction with the ion are very important. In unshielded, field electrolytes, the solvent merely acts as a dielectric medium. In Fig. 1 the shaded zones indicate, within the uncertainty of the dielectric constant data,[19] the Born[20,21] electrostatic heat capacities for the two electrolytes, which have been obtained using the higher-temperature data to fix the radius terms; the agreement with simple electrostatic behaviour, above ~100 °C, is good. The temperature region where $(\mathrm{d}\bar{C}^{\circ}_{p2}/\mathrm{d}T) = 0$ is a good working definition of the transition region between the two types of behaviour, which is between 40° and 70 °C for many species (Fig. 2). Table 1 summarizes

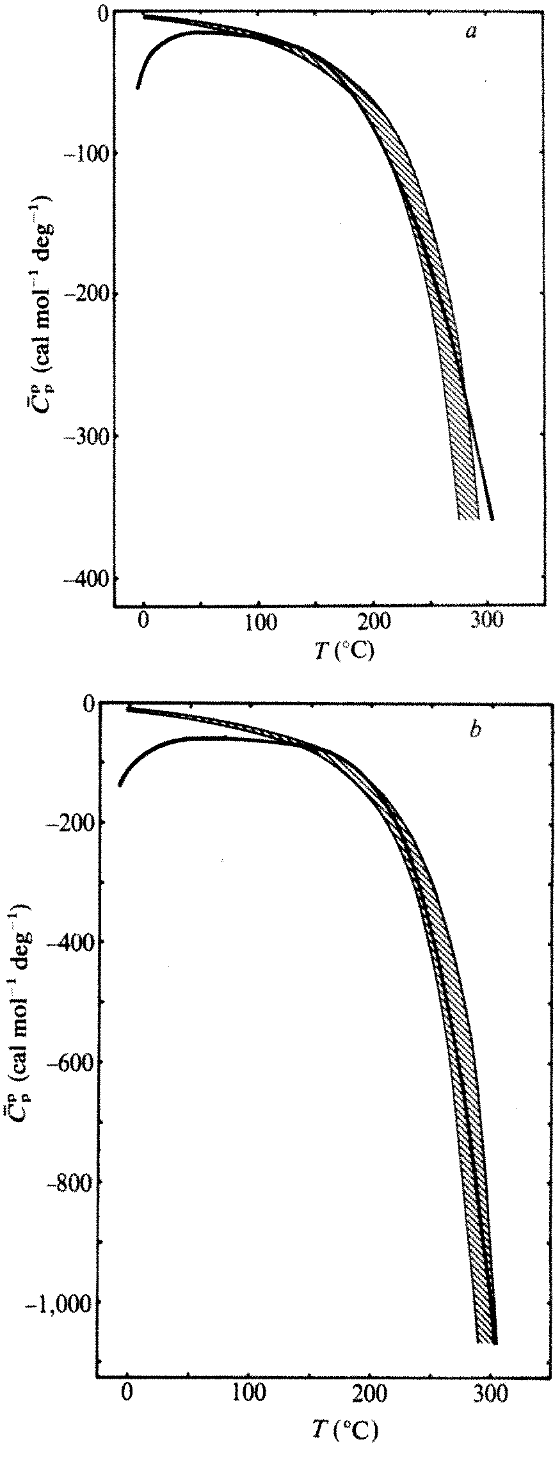


Fig. 2 Standard-state partial molal heat capacities for selected electrolytes[29,30], 0–100 °C.

some of the expected differences in behaviour between 'field' and 'structure' electrolytes.

There are other data which support these arguments. Twice differentiated free energies from precise equilibrium measurements in dilute solutions can give approximate values of $\bar{C}^{\circ}_{p2}$ at higher temperatures. The resulting large negative values, in cal mol^{-1} deg^{-1} at 300 °C, are -214 for $H^+ + OH^-$ from $K_w(H_2O)$ (ref. 22), -210 for $NH_4^+ + OH^-$ from K_d (ref. 23) and the aqueous solubility[24] of ammonia, and -447 for $2H^+ + CO_3^{-2}$ from the aqueous dissociation[25,26] and the solubility[27] of CO_2. Furthermore, the large, negative partial molal volumes ($\sim -10^4$ cm^3)[28] obtained for the ionic dissociation of NaCl(aq) and HCl(aq) above 360 °C are believed to be consistent with our model. We are not aware of extensive heat capacity data in other solvent systems.

This paper is the result of work supported over previous years by the NSF and presently by the Electric Power Research Institute.

Received 26 November 1980; accepted 16 April 1981.

1. Criss, C. M. & Cobble, J. W. *J. Am. chem. Soc.* **83,** 3223–3228 (1961).
2. Ellis, A. J. & McFadden, I. M. *Geochim. cosmochim. Acta* **36,** 413–426 (1972).
3. Kelley, K. K. *U.S. Bur. Mines Bull.* 584 (1960).
4. Gardner, W. L., Mitchell, R. E. & Cobble, J. W. *J. phys. Chem.* **73,** 2025–2032 (1969).
5. Wood, R. W. & Smith-Magowan, D. *Am. Chem. Soc. Ser.* No. 133, 569–580 (1980).
6. Kavanau, J. Lee *Water and Solute–Water Interactions* (Holden-Day, San Francisco, 1964).
7. Eisenberg, D. & Kauzmann, W. *The Structure and Properties of Water* (Oxford University Press, 1969).
8. Narten, A. H. & Levy, H. A. *Science* **165,** 447–454 (1969).
9. Frank, H. S. *Science* **169,** 635–641 (1970).
10. Angell, C. A. *J. phys. Chem.* **75,** 3698–3705 (1971).
11. Hagler, A. T., Scheraga, H. A. & Nemethy, G. *J. phys. Chem.* **76,** 3229–3243 (1972).
12. Benson, S. W. *J. Am. chem. Soc.* **100,** 5640–5644 (1978).
13. Frank, H. S. & Wen, W. Y. *Discuss. Faraday Soc.* 133–140 (1957).
14. Conway, B. E., Desnoyers, J. E. & Smith, A. C. *Phil. Trans. R. Soc.* **A256,** 389–437 (1964).
15. Rosseinsky, D. R. *Chem. Rev.* **65,** 467–490 (1965).
16. Bockris, J. O'M. & Saluja, P. P. S. *J. phys. Chem.* **76,** 2298–2310 (1972).
17. Haggis, G. H., Hasted, J. B. & Buchanan, T. J. *J. chem. Phys.* **20,** 1452–1465 (1952).
18. Cobble, J. W. *Science* **152,** 1479–1485 (1966).
19. Bradley, D. H. & Pitzer, K. S. *J. phys. Chem.* **83,** 1599–1603 (1979).
20. Born, M. *Z. Phys.* **1,** 45–48 (1920).
21. Cobble, J. W. & Murray, R. C. Jr *Trans. Faraday Soc.* **64,** 144–149 (1978).
22. Sweeton, F. H., Mesmer, R. E. & Baes, C. F. Jr *J. Solut. Chem.* **3,** 191–214 (1974).
23. Hitch, B. F. & Mesmer, R. E. *J. Solut. Chem.* **5,** 667–679 (1976).
24. Jones, M. E. *J. phys. Chem.* **67,** 1113–1115 (1963).
25. Read, A. J. *J. Solut. Chem.* **4,** 53–70 (1975).
26. Helgeson, H. A. *J. phys. Chem.* **71,** 3121–3136 (1967).
27. Ellis, A. J. & Golding, R. M. *Am. J. Sci.* **261,** 47–60 (1963).
28. Pearson, D., Copeland, C. S. & Benson, S. W. *J. Am. chem. Soc.* **85,** 1047–1049 (1963).
29. Gardner, W. L., Jekel, E. C. & Cobble, J. W. *J. phys. Chem.* **73,** 2017–2020 (1969).
30. Cobble, J. W., Stephens, H. P., McKinnon, I. R. & Westrum, E. F. Jr *Inorg. Chem.* **11,** 1669–1674 (1972).

Blast dynamics at Mount St Helens on 18 May 1980

Susan Werner Kieffer

US Geological Survey, Flagstaff, Arizona 86001, USA

At 8.32 a.m. on 18 May 1980, failure of the upper part of the north slope of Mount St Helens triggered a lateral eruption ('the blast') that devastated the conifer forests in a sector covering ~500 km^2 north of the volcano. I present here a steady flow model for the blast dynamics and propose that through much of the devastated area the blast was a supersonic flow of a complex multiphase (solid, liquid, vapour) mixture. The shape of the blast zone; pressure, temperature, velocity (Mach number) and density distributions within the flow; positions of weak and strong internal shocks; and mass flux, energy flux, and total energy are calculated. The shape of blast zone was determined by the initial areal expansion from the reservoir, by internal expansion and compression waves (including shocks), and by the density of the expanding mixture. The pressure within the flow dropped rapidly away from the source of the blast until, at a distance of ~11 km, the flow became underpressured relative to the surrounding atmosphere. Weak shocks within the flow subparallel to the east and west margins coalesced at about this distance into a strong Mach disk shock, across which the flow velocities would have dropped from supersonic to subsonic as the pressure rose back towards ambient. The positions of the shocks may be reflected in differences in the patterns of felled trees. At the limits of the devastated area, the temperature had dropped only 20% from the reservoir temperature because the entrained solids thermally buffered the flow (the dynamic and thermodynamic effects of the admixture of the surrounding atmosphere and the uprooted forest and soils into the flow are not considered). The density of the flow decreased with distance until, at the limits of the blast zone, 20–25 km from the volcano, the density became comparable with that of the surrounding (dirty) atmosphere and the flow became buoyant and ramped up into the atmosphere. According to the model, the mass flux per unit area at the source was 0.6 × 10^4 g s^{-1} cm^{-2} and the energy flux per unit area was 2.5 MW cm^{-2}. From the measured total ejected mass, 0.25 × 10^{15} g, the total energy released during the eruption was 10^{24} erg or 24 megatons. The model, triggering of the eruption and the transition from unsteady to steady flow, and applications to eyewitness observations and atmospheric effects are discussed in ref. 1.

The eruption began through a vent centred between ~1,900 and 2,500 m elevation ~1 km north of the old summit of the mountain. The 'devastated area' (the area in which trees have been destroyed) is a sector generally extending to the north of an east–west line through the vent formed by the landslide that triggered the event (Fig. 1). The devastated area is surrounded by a zone, the 'singed zone', in which trees are still standing, but on which remaining needles were singed.

Preliminary mapping of the direction of fall of the trees shows two major irregularly shaped zones: (1) an inner zone, the 'direct blast zone', in which the flow was approximately radial from the volcano and was relatively undeflected, in plan view, by even large topographic features; and (2) an outer zone, 'the channelized blast zone', in which the flow followed or was deflected, in plan view, by the local topography (Fig. 1). Streamlines of the flow were thus not simple radii from the source.

The reservoir from which the material erupted is assumed to have been a disk-shaped volume oriented vertically on the scarp formed on the north side of the mountain by the landslide. It is taken to be just large enough to contain the mass ejected by the blast, 0.25 × 10^{15} g (ref. 2). At an assumed pre-eruption bulk density of 2.0 g cm^{-3}, the reservoir would have had a volume of 0.12 km^3. The frontal surface of this disk (the vent) is assumed to have been a rectangle 1 km in east–west dimension and 0.25 km in height; the reservoir would therefore have been 0.5 km deep. The acceleration of the blast material from rest to the vent velocity discussed below is assumed to have taken place within the reservoir[1].

The erupted material is characterized in this model by an initial pressure, P_0, an initial temperature, T_0, and an initial mass ratio, m, of solid plus liquid phases to the vapour phase. For an adiabatic, quasi-isentropic expansion, a mixture of vapour and fine-grained solid (or inert liquid) particles can be modelled as a thermally-equilibrated pseudo-gas and can be characterized by perfect gas equations of state[3]:

$$PV = R_b T \tag{1a}$$

$$PV^{\gamma_b} = \text{constant} \tag{1b}$$

where P is the pressure, V is the volume, γ_b is the isentropic exponent, and R_b is the gas constant of the blast mixture. Because of the fine-grained nature of the entrained solid particles, thermal transfer of heat to the gas phase from the solids is important; is such a case, the values of γ_b and R_b are

$$\gamma_b = \frac{C_{P\text{g}} + mC_\text{s}}{C_{V\text{g}} + mC_\text{s}} \tag{2a}$$

$$R_b = R_g/(1+m) \tag{2b}$$

0028-0836/81/250568—03$01.00

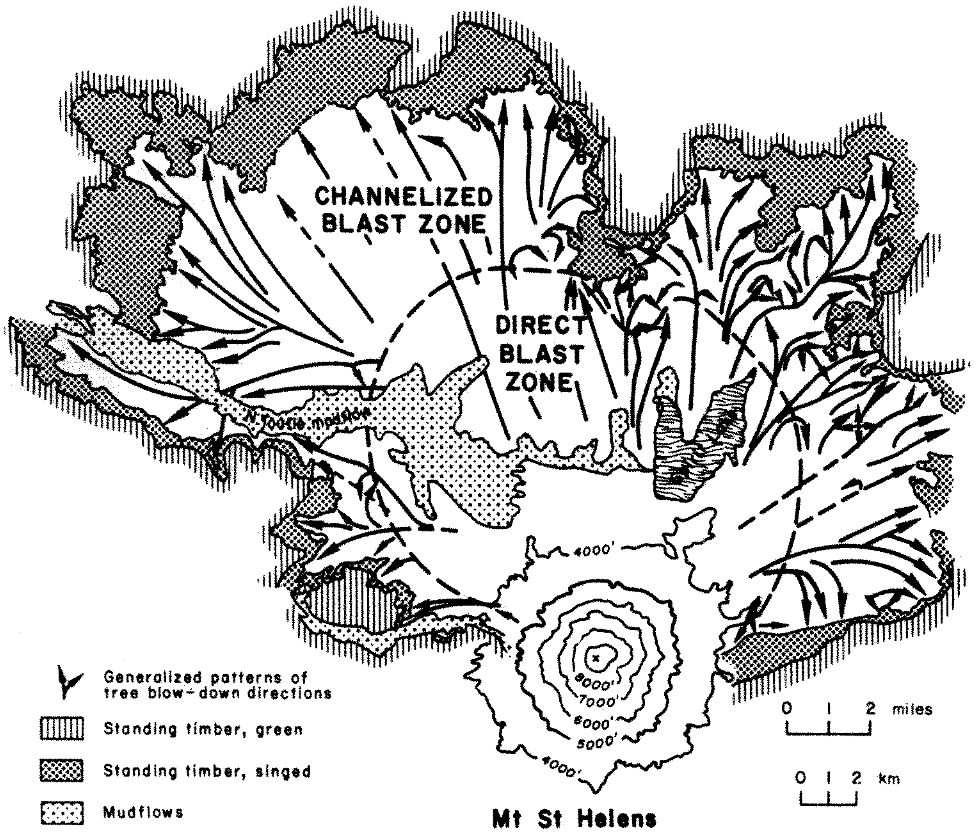


Fig. 1 The devastated area, and flow streamlines (arrowed vectors) as indicated by the alignment of fallen trees. The landslide at 8.32 a.m. initially broke ~1km north of the old summit and formed the near-vertical vent from which the blast emanated. Relative to this vent position, the margin of the flow on the volcano bends south of due east and due west and then sweeps north-east and north-west as it diverges from the summit of the mountain. The singed zone is narrow compared with the size of the devastated area and streamlines of the flow are generally perpendicular to the margin. The boundary between the direct and channelized blast zones is shown by the short broken lines.

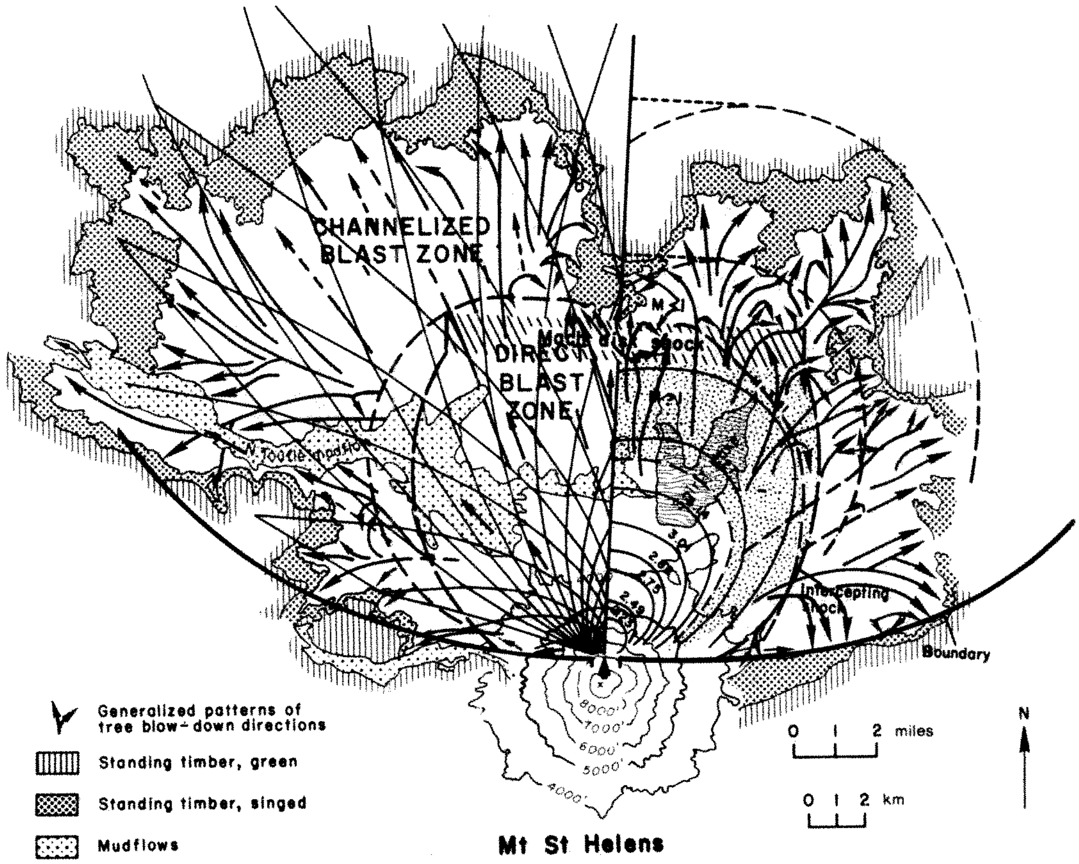


Fig. 2 The model flow field superimposed on the map of Fig. 1. The vent width has been taken as 1 km, has been placed at the 8,000-ft contour level north of the old summit of the mountain, and has been orientated 5° east of due north. As the model is symmetric about the axis of the vent it can be split into two halves. On the left, the mathematical characteristics of the solution are shown as thin lines radiating from the corner of the vent and from the mirror plane on the axis of symmetry. The arrows show the directions of the flow and would be the general direction of felled trees. The heavy line indicates the boundary of the flow, a 0.87-bar isobar in the model. The intercepting shock formed by reflection of the expansion waves from this boundary is shown as the long-dashed line connecting the Mach disk shock. Lines of constant Mach number and, therefore, constant pressure, temperature and density are shown. The approximate location of the Mach disk shock is shown by the sloping pattern 8 km across. Because estimates of flow density are needed for hazards zone prediction, the analytical solution is extrapolated into the subsonic region beyond the Mach disk shock (dashed contours), and is assumed to be about valid for density if the shocks are not strong. Each contour line is labelled by the value of the Mach number, *M*. The values of M, P/P_0, T/T_0 and ρ/ρ_0 are, in the supersonic region, 2.23, 0.087, 0.91, 0.095; 2.49, 0.047, 0.89, 0.053; 2.75, 0.025, 0.87, 0.029; 2.88, 0.018, 0.86, 0.021; 3.01, 0.013, 0.85, 0.016; and 3.14, 0.009, 0.83, 0.011. Extrapolated into the subsonic region: 3.27, 0.006, 0.82, 0.007; 3.41, 0.004, 0.81, 0.006; 3.54, 0.003, 0.80, 0.004; 3.68, 0.002, 0.785, 0.002.

where C_{P_g} and C_{V_g} are, respectively, the specific heats at constant pressure and volume of the gas, and C_s is the specific heat of the solid phase (for which C_P and C_V are assumed to be approximately equal). Although the model can easily be scaled with dimensional, pressure and temperature ratios, a model case for a single set of plausible initial conditions at Mt St Helens is presented here. For this case, it was assumed that 10% by volume of the flank of the volcano that became the blast material was filled with liquid water that vaporized to give the gas during the eruption; for an average rock density of 2.5 g cm^{-3} and liquid water density of 1 g cm^{-3}, this assumption gives $m = 25$, $\gamma = 1.04$, and $R_b = 17.8 \times 10^4$ erg C^{-1} mol^{-1}. The initial source pressure is assumed to have been 12.5 Mpa, appropriate to 650 m of overburden, and the initial temperature is assumed to have been 600 K, the saturation temperature of pure H_2O at 12.5 MPa. The initial density of a mixture of rock (assumed to have vesiculated or cracked to an average density of 2.0 g cm^{-3}) and steam (at a density of 0.074 g cm^{-3}) at 12.5 MPa pressure would have been 1.0 g cm^{-3}.

Because the reservoir was at such high pressure compared with the ambient atmospheric pressure, the flow would have choked at the vent and exited at the choked or sonic flow velocity. For steady flow, this velocity is

$$u_s{}^* = \left(\frac{2}{\gamma_b + 1}\right)^{1/2} c_{0b} \qquad (3)$$

where c_{0b} is the speed of sound of the blast mixture at rest:

$$c_{0b} = (\gamma_b R_b T)^{1/2} \qquad (4)$$

For the mixture postulated, c_{0b} is 105 m s^{-1} and the choked velocity is 104 m s^{-1}. The pressure at the vent would have decreased to $P^* = 0.60\ P_0$ (7.5 MPa), the temperature to $T^* = 0.98\ T_0$ (588 K), and the density to $\rho^* = 0.61\rho_0$ (0.61 g cm^{-3}). The mass flux, controlled at the vent by choking, was

$$\dot{m}^* = \rho^* u^* A^* \qquad (5)$$

where ρ^*, u^* and A^* are, respectively, the density, velocity and area at sonic conditions at the vent. With an average temperature 325 K above ambient, a specific heat of 0.8×10^7 erg g^{-1} K^{-1} for the entrained solid or liquid phases, 4.2×10^7 erg g^{-1} K^{-1} for the steam, and 2.3×10^{10} erg g^{-1} latent heat of vaporization of the steam, the thermal energy flux or power unit area was 2.5×10^{13} erg s^{-1} cm^{-2}, or 2.5 MW cm^{-2}. From the measured mass ejected and the assumed initial temperature of 600 K, the total thermal energy released on cooling to ambient temperature (273 K) was 10^{24} erg or 24 megatons, of which 7 megatons were released during the propagation of the blast through the devastated area, and 17 megatons during penecontemporaneous condensation of the steam and cooling of the condensed water and entrained solids.

After leaving the vent at sonic velocity, the material would have expanded supersonically in all directions into the lower-pressure atmosphere. In the absence of gravity, friction, phase changes, and addition or loss of mass from a flowing fluid, the steady-state equations of motion for an expanding fluid are:

$$\bar{\nabla} \cdot (\rho \bar{u}) = 0 \qquad \text{(continuity} \quad (6a)$$

$$\bar{\nabla} \times \bar{u} = 0 \qquad \text{(irrotationality)} \quad (6b)$$

$$\bar{\nabla}\left(\frac{\bar{u}^2}{2}\right) + \frac{1}{\rho}\bar{\nabla} P = 0 \qquad \text{(momentum)} \quad (6c)$$

Here ρ is the density, $\bar{u}$ is the vector velocity, P is the pressure, and $\bar{\nabla}$ is the spatial differential operator. These equations, supplemented by the equation of state of the fluid and the boundary condition that the edge of the flow is everywhere at atmospheric pressure, define the steady state condition of the fluid. They were solved by the method of characteristics within the region where shocks do not exist; various scaling relations adopted from the aeronautics literature[4] were used to estimate the position and size of shocks, and the shape of the flow boundary outside the regions of shocks.

The flow model for the boundary and reservoir conditions assumed above for Mt St Helens is shown in Fig. 2. The characteristics of the solution can be thought of as expansion waves across which the flow accelerated as the pressure, density and temperature dropped. For the model conditions, the flow would have expanded initially through an angle of 96° near the vent. At the boundary of the flow, the pressure was reduced to the local atmospheric pressure and the initial expansion waves were reflected back into the flow as compression waves that coalesced into the intercepting shocks shown near the margins of the flow. These reflections deflected the flow from its initial angle of 96° towards ~60° at a distance of 10 vent diameters out from the mountain.

As the flow diverged, the pressure dropped rapidly away from the vent. Isobars show that the pressure dropped below ambient outside a zone ~6 vent diameters from the vent (the negative excess pressure zone is shown as a stippled area in Fig. 2). This low-pressure core was stable within the flow because the compression waves could not be reflected back into the flow quickly enough to prevent overexpansion. The Mach disk shock is a consequence of the overexpansion and, across this shock, the pressure rose back towards ambient, and the flow velocities decreased from supersonic to subsonic. I proposed that the position of the Mach disk shock and the lateral intercepting shocks coincides with the boundary between the direct and channelized blast zones, and that the transition from supersonic to subsonic flow across this boundary accounts for differences in tree blowdown patterns.

As the pressure within the flow decreased, the density likewise decreased rapidly (although temperatures were remarkably uniform across the flow because of the buffering effect of the entrained solid and liquid phases which acted as a heat source for the expanding vapour). The density dropped from a value of 1.0 g cm^{-3} in the reservoir to a value very near that of the atmosphere (which is assumed to have had 2–3 times its normal density because of dust-ladening during the unsteady development of the flow) at ~20–25 km distance from the vent. (It is assumed that the densities given by the solution interior to the shocks can be extrapolated across the shocks, a reasonable approximation if the shocks are not too strong.) Thus, as the density of the flow became equal to that of the (dirty) atmosphere, the flow became buoyant and simply lifted from the ground and ramped up into the atmosphere. The singed zone then is the area across which the flow was still close enough to the ground to cause thermal effects as it rose.

Numerical uncertainties for the calculated variables would be meaningless in view of the uncertainties in physical principles that we still face in interpreting the blast flow dynamics. The important reservoir parameters suggested by this model are the initial pressure (because it determines the amount of lateral divergence as the flow leaves the vent), the mass ratio, m, of solid-to-vapour components (because it determines the isentropic exponent, γ, and thus the thermodynamic history of the flow), and the vent diameter (because the dimensions of the devastated area scale with this dimension). From the model, however, it seems that, given reasonable *a priori* knowledge of the location of a potential vent and possible initial reservoir conditions, a hazards zone for blast-like events could be estimated by a relatively simple calculation. Models such as this and, as available, computer models that can calculate flow properties across internal shock waves, will be necessary for delineation of blast hazard zones because little evidence of the blast events is preserved in the geological record.

Received 9 February; accepted 25 March 1981.

1. Kieffer, S. W. *U.S. geol. Surv. Prof. Pap.* 1250 (in the press).
2. Moore, J. G. & Albee, W. C. *U.S. geol. Surv. Prof. Pap.* 1250 (in the press).
3. Wallis, G. B. *One-dimensional Two-phase Flow* (McGraw Hill, New York, 1972).
4. JANNAF (Joint Army, Navy, NASA, Air Force) *Handbook of Rocket Exhaust Plume Technology*, Ch. 2 (CPIA Publ. 263, 1975).

A model for the origin of the alkaline complexes of Egypt

Phil de Gruyter & Thomas A. Vogel

Department of Geology, Michigan State University, East Lansing, Michigan 48824, USA

The alkaline complexes in the Nubian Shield of Egypt range in age from 554 to 89 Myr BP (refs 1–3) and, where data are available, have similar and primitive initial strontium ratios of ^{87}Sr/^{86}Sr (ref. 1). These intrusive complexes are similar with respect to petrography, major element chemistry, rare earth element distribution and other trace element distribution[1,4–7]; they are intimately associated with major lineaments in the Nubian Shield (Fig. 1)[8]. It is likely that all of them had a similar origin, not directly related to rifting and/or doming. We suggest here that the origin of these complexes is due to alkaline melts having been formed in the asthenosphere by shear heating, caused by changes in plate motion. These melts were emplaced along reactivated Pan-African fractures or pre-existing zones of weakness.

The major structural features of the Nubian Shield were formed during the Pan-African compression–accretion process (1,000–550 Myr BP)[9–11]. This NE–SW oblique compression (see Fig. 4) produced both NE–SW and NW–SE zones of weakness[9,12]. Most of the alkaline complexes of Egypt are associated with these lineaments.

The oldest known[2] alkaline intrusion in the Eastern Desert of Egypt is Wadi Dib, 554 Myr BP, emplaced during the last stages of the Pan-African event[9]. This complex is associated with the north-west trending Najd transcurrent fault system (Fig. 1) which has ~240 km of lateral displacement[13] and developed at about the same time span (580–530 Myr BP)[14] as the emplacement of the alkaline complex of Wadi Dib.

Zargat Naam (404 Myr BP)[2] may be associated with a north-west lineament and Tarbite North is located near the intersection of two lineaments (Fig. 1). El Gezira is dated[2] at 229 Myr BP and is associated with the same major north-east trending lineament as El Naga, El Mishbeh, Nigrub El Tahiani and Nigrub El Fogani. The ages[2] of these complexes are 145, 142, 140, 139, 132 Myr BP respectively. El Kahfa and Abu Khrug (89 and 91 Myr BP)[1] are associated with the same set of north-westerly trending lineaments (Fig. 1).

Although the alkaline complexes of the Egyptian Eastern Desert have similar overall chemical trends, there is a distinct change in the degree of alkalinity and silica saturation with age of the complexes[4]. The older complexes are generally silica oversaturated and relatively low in alkalis, whereas the younger complexes are dominated by silica-undersaturated rock types and are rich in alkalis (Fig. 2). This change in alkalinity may be a direct result of the cooling of the lithosphere after formation of the Pan-African thermal event. Higher temperatures after the Pan-African event would be conducive to larger degrees of melting at shallow levels, thus producing less alkaline magmas. As the lithosphere cooled, smaller degrees of melting occurred at greater depths, and more alkaline melts would be produced. Thus the chemical trends observed in the Egyptian alkaline complexes are consistent with decreasing thermal gradients in the Phanerozoic.

The six periods of alkaline magmatism in Egypt (554, 404, 351, 229, 145–132, 91–89 Myr BP) are all synchronous with known changes of plate motion. Changes in plate motion can be inferred from: (1) abrupt changes in global sea-level curves[15]; (2) the polar wandering curve for Africa since Pan-African time; and (3) the motion of adjacent plates during the past 180 Myr. Table 1 summarizes these data.

In comparison with tholeiitic magmatism, it is widely accepted that alkaline magmatism has a relatively deep-seated origin[16–19] and may be associated with relatively small degrees of partial melting of an enriched mantle source[20–22]. These melts ascend preferentially along zones of weakness.

In light of recent thermo-mechanical modelling[23–30], a shear heating model for alkaline melts is reasonable for the production of alkaline magmas in areas of relatively low heat flow. Briefly, thermo-mechanical modelling indicates that an increase in plate velocity or a rapid change in direction of motion, causes an increase of the shear stress on the base of the lithosphere. This leads to temperature increases, viscosity decreases, and further plate stress and velocity changes. In these conditions the asthenosphere, the zone of decoupling between the rigid lithosphere and the viscous mesosphere, may be the locus for metasomatic fluids. In this region melts may be generated in sufficient quantities to become a significant source of alkaline magmatism[31–34]. This melt production probably occurs at different levels[36] in a potentially unstable, thermally-active 'boundary zone' of the lower lithosphere–asthenosphere[35] (Fig. 3). It is suggested here that alkaline complexes of the Egyptian Eastern Desert area may be directly related to acceleration in plate motion. In this model, melting results from the internal dissipation of heat produced by this motion. Factors which are likely to have a bearing on the incidence of alkaline magmatism are: the amount of thermal disequilibrium produced by localization of shear heating; the ease of ascent of the melts (presence of fractures); and the possible involvement of non-alkaline melts generated at shallower depths. This approach to alkaline magmatism takes into account several features common to many alkaline complexes, including those in Egypt. They are: (1) Alkaline magmatism is unrestricted in its geographic and tectonic setting. It occurs in every tectonic environment in the oceans or continents[37–52], and is often independent of any rifting and/or doming events. (2) Alkaline complexes are emplaced in terrains that have relatively low thermal gradients[38,53]. (3) They

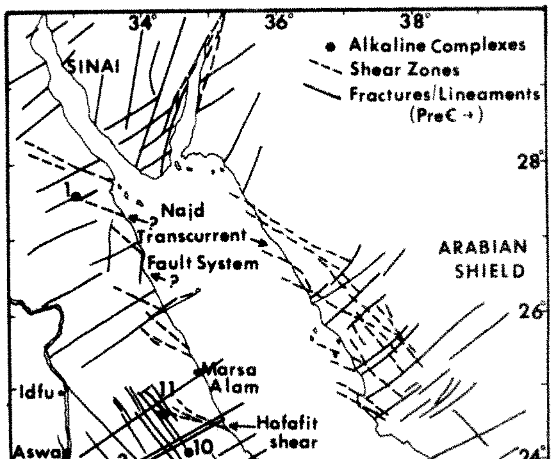


Fig. 1 1, Wadi Dib 554 Myr BP; 2, Zargat Naam 404 Myr BP; 3, Tarbtie (North) 351 Myr BP; 4, El Gezira 229 Myr BP; 5, El Naga 145 Myr BP; 6, El Mishbeh 142 Myr BP; 7, Nigrub El Tahtani 140 Myr BP; 8, Nigrub El Fogani 139 Myr BP; 9, Mansouri 132 Myr BP; 10, El Kahfa 91 Myr BP; 11, Abu Khrug 89 Myr BP. Dates from refs 2, 3. Map and tectonic lineaments after ref. 8.

0028–0836/81/250571—04$01.00

Table 1 Changes in plate motion

Periods of tectonic activity	Periods of rapid polar wandering (Africa) (Myr)[70]	Rapid changes in sea level[15] (Myr)	Periods of relative motion between Africa and Europe (Myr)	Periods of alkaline magmatism in Egypt[1–3] (Myr)
Pan-African event of Eastern Desert, Egypt 1,000–550 Myr BP[9–11]. Large-scale transcurrent faulting (Najd fault system, active 580–530 Myr BP towards the end of Pan African event or Hijaz Orogenic Cycle)[13,14,87]	680–660	480 410–390		554 404
Initial break-up of Pangea[71] leading to extension of North America–Africa/South America	400–220	350 330–320		351
Overall compression and subduction of Palaeo-Tethys north of Afro-Arabia; possible back-arc-type spreading of Cimmerian land mass from NE portion of Gondwana (including Afro-Arabia) and formation of Neo-Tethys[72]		280 220–185		229, 223, 221
Initial rifting of the South Atlantic: Africa–South America		140	165–80	145, 142, 140
and Benue Trough extension			Mid- to late Cretaceous	139, 132
Corresponding compression in Afro-Arabia and Tethyan domain strike-slip faulting, closely followed by extension between Iran and Arabia[73,74] and the possible separation of the Anatolian plate from Afro-Arabia[75]				
Major shift in spreading of the South Atlantic and associated second major episode of alkaline magmatism[56,57,60,61] ~80–90 Myr BP Large-scale strike-slip faulting in Afro-Arabia[5,76–81] intense compression in Iran and Turkey with associated ophiolites[73,74,82–84], emplacement of Oman ophiolites in NE Arabia[85]	110–40	95 60	80–40 Late Cretaceous Lower Eocene	91,89
Red Sea doming and extension	40–0 Lower Eocene to Recent	30		Recent alkali magatism associated with Red Sea opening

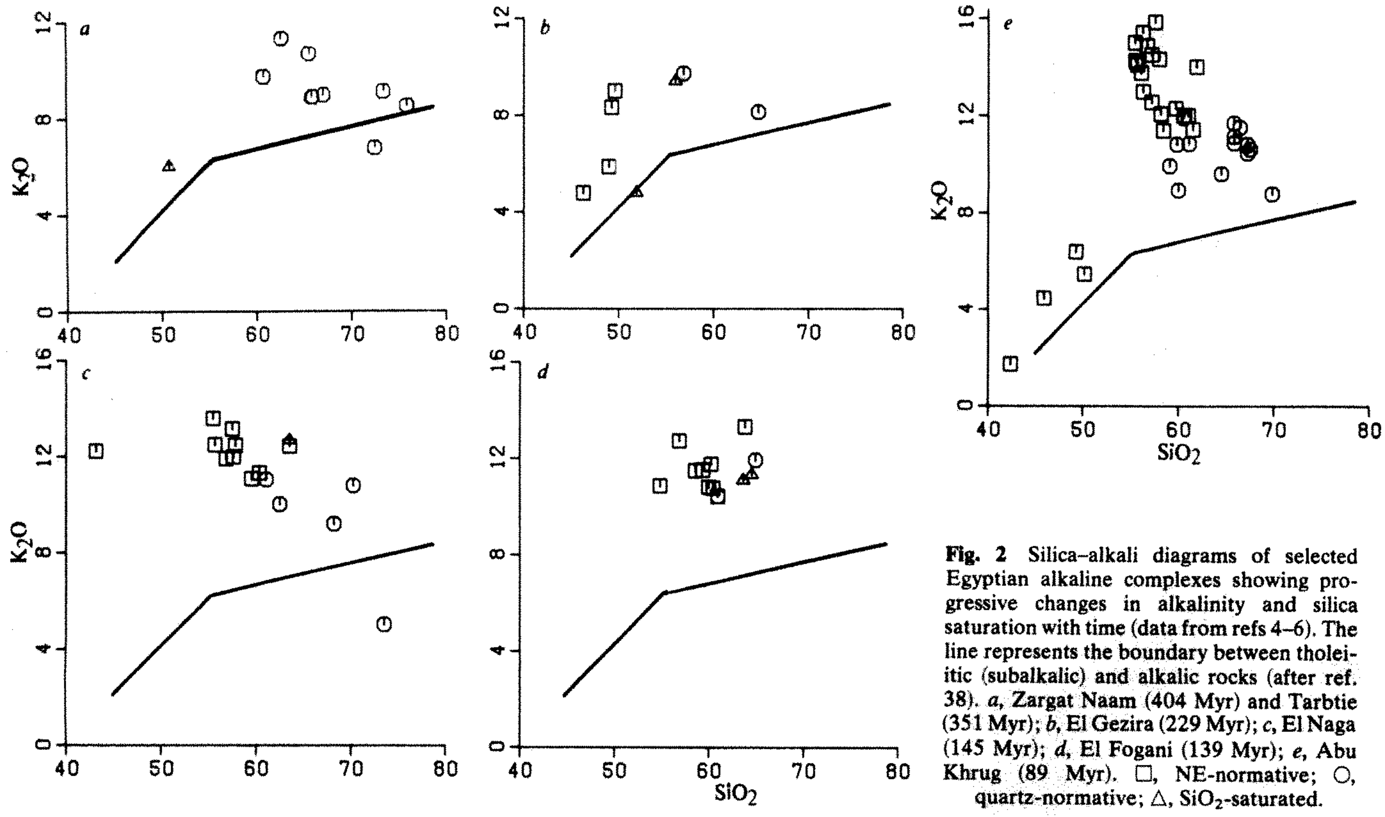


Fig. 2 Silica–alkali diagrams of selected Egyptian alkaline complexes showing progressive changes in alkalinity and silica saturation with time (data from refs 4–6). The line represents the boundary between tholeiitic (subalkalic) and alkalic rocks (after ref. 38). *a*, Zargat Naam (404 Myr) and Tarbtie (351 Myr); *b*, El Gezira (229 Myr); *c*, El Naga (145 Myr); *d*, El Fogani (139 Myr); *e*, Abu Khrug (89 Myr). □, NE-normative; ○, quartz-normative; △, SiO_2-saturated.

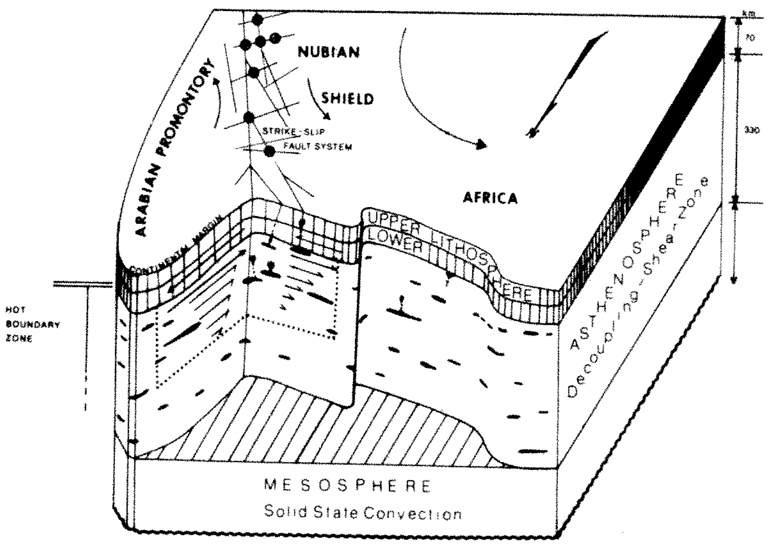


Fig. 3 Alkaline melt production through shear heating in the asthenosphere in response to major shifts in the motion of the African plate and their ascent along the reactivated strike-slip fault system (shear zones) between the Arabian Promontory and the Nubian shield. The generation of small melt fractions through internal heat dissipation is illustrated in the zone of decoupling between the lithosphere and mesosphere in response to major shifts in lithospheric motion. Similar small melt fractions may be produced in response to sustained opposing plate motion (for example, along a transform fault system). Dotted area outlines a general region of melt release and concentration of shear dissipation. No scale intended. ●, Alkaline magmatism.

have similar chemical trends and are enriched in some trace elements particularly the large ionic lithophile elements[4,20,37,38,53,55]. (4) Egyptian alkaline complexes, along with other alkaline complexes, tend to occur during certain ages[56–59], and these ages generally correspond to known accelerations in plate motion[39,40,44,56,57,60–63]. (5) Many alkaline complexes are associated with long-lived zones of crustal weakness[57–68].

The coincidence of alkaline magmatism in widely separated areas is well illustrated in the case of the Egyptian and Atlantic alkaline provinces. The 224 Myr episode of alkaline magmatism in Egypt coincides with the first episode of alkaline magmatism in the White Mountain Magma Series[62] along the crustal lineaments within the (still) adjoining North American continent. The 140 and 90 Myr episodes of alkaline magmatism in Egypt coincide with the two major episodes of alkaline magmatism occurring in the areas surrounding the South Atlantic. Alkaline magmatism around 140 Myr in Brazil, Uruguay, Bolivia, and western and southern Africa has been related to the initial rifting

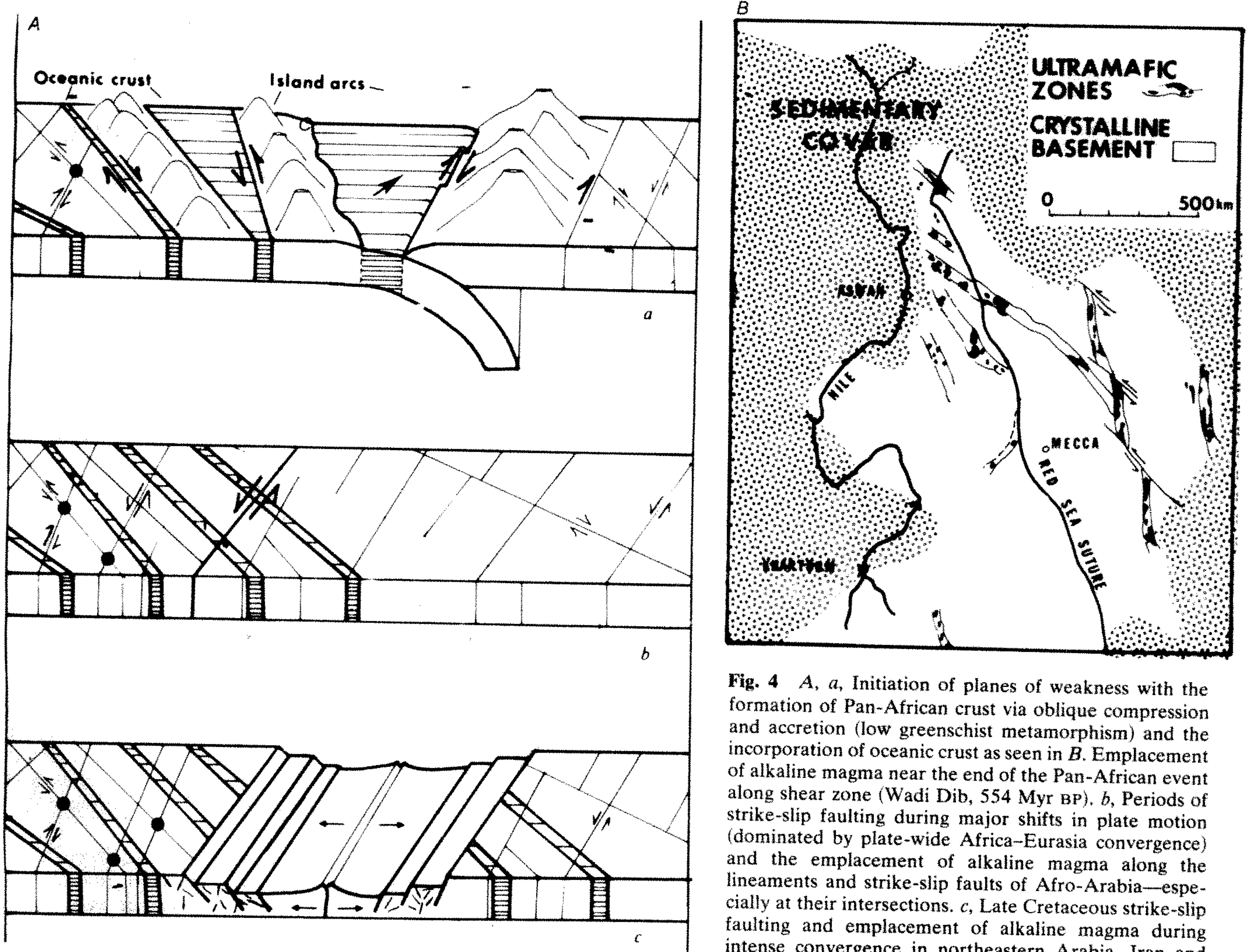


Fig. 4 *A*, *a*, Initiation of planes of weakness with the formation of Pan-African crust via oblique compression and accretion (low greenschist metamorphism) and the incorporation of oceanic crust as seen in *B*. Emplacement of alkaline magma near the end of the Pan-African event along shear zone (Wadi Dib, 554 Myr BP). *b*, Periods of strike-slip faulting during major shifts in plate motion (dominated by plate-wide Africa–Eurasia convergence) and the emplacement of alkaline magma along the lineaments and strike-slip faults of Afro-Arabia—especially at their intersections. *c*, Late Cretaceous strike-slip faulting and emplacement of alkaline magma during intense convergence in northeastern Arabia, Iran and Anatolia followed by mantle upwelling and rifting along intensely weakened shear zones (Red Sea). *B*, Possible Pan-African ophiolites within Afro-Arabian shield with palinspastic reconstruction of Red Sea area, after ref. 12. ●, Alkaline magmatism.

of Africa from South America[56–59,64]. The alkaline magmatism around 80–90 Myr has been correlated with a major shift in the spreading of the South Atlantic during the late Cretaceous[56,57,60,63]. Considering the reciprocal plate motion involved in the opening of the South Atlantic and the closing of the Tethyan ocean basins[65–67], the synchronous emplacement of these alkaline melts is not likely to be the result of unrelated mantle disturbances. We suggest that shear-heating in the asthenosphere due to these large changes in plate motion is a more likely explanation. Alternatively, these alkaline melts may be due to worldwide, synchronous, mantle disturbances[69].

The alkaline complexes of Egypt have been emplaced over a large time span (450 Myr) and seem to be associated with reactivated Precambrian lineaments. The ages of these alkaline complexes appear to be synchronous with changes in plate motion. Our suggestion for the origin of these alkaline complexes is that heat, necessary for melting, is produced through shear melting during acceleration of plate motion (Fig. 3) and that they are emplaced along pre-existing fractures or zones of weakness[86]. This model may have general relevance to the origin of other alkaline magmas.

This research has been supported by NSF grant OIP 75-07943 administered by the University of South Carolina. We thank W. H. Kanes and E. M. El Shazly for continued interest and support, also Ahmed Abdallah for field assistance.

Received 20 January; accepted 2 April 1981.

1. Lutz, T. M. thesis, Univ. Pennsylvania (1979).
2. Serencsits, C. McC., El Ramly, M. F., Faul, H., Foland, K. A. & Hussein, A. A. *Geol. Surv. Egypt* (in the press).
3. Hashad, A. H. & El Reedy, M. W. M. *5th Conf. on African Geology Abst.*, Cairo, 44–45 (1979).
4. deGruyter, P., & Vogel, T. A. (in preparation).
5. El Ramly, M. F., Dereniuk, N. E., Budanov, V. I., Armonious, L. K. & Hayek, G. G. *Geol. Surv. Egypt Pap.* No. 51 (1969).
6. El Ramly, M. F., Armanious, L. K., Budanov, V. I. & Derniuk, N. E. *Geol. Surv. Egypt Pap.* No. 52 (1969).
7. El Ramly, M. F., Budanov, V. I. & Hussein, A. A. A. *Geol. Surv. Egypt Pap.* No. 53 (1971).
8. Garson, M. S. & Krs, M. *Bull. geol. Soc. Am.* **87**, 169–181 (1976).
9. Gass, I. G. *J. geol. Soc. Lond.* **134**, 129–138 (1977).
10. Engel, A. E. J., Dixon, T. H. & Stern, R. *5th Conf. on African Geology Abstr.*, Cairo, 35–36 (1979).
11. Engel, A. E. J., Dixon, T. H. & Stern, R. J. *Bull. geol. Soc. Am.* **91**, 699–706 (1980).
12. Neary, C. R., Gass, I. G. & Cavanaugh, B. J. *Bull. geol. Soc. Am.* **87**, 1501–1512 (1976).
13. Moore, J. Mc. *J. geol. Soc. Lond.* **136**, 441–454 (1979).
14. Fleck, R. J. *et al. Bull. geol. Soc. Am.* **87**, 9–21 (1976).
15. Vail, P. R., Mitchum, R. M. Jr & Thompson, S. III in *Seismic Stratigraphy—Applications to Hydrocarbon Exploration* Vol. 26, 83–98 (AAPG MEM, 1977).
16. Green, D. H. & Ringwood, A. E. *Contr. Miner. Petrol.* **15**, 103–190 (1967).
17. Bultitude, R. J. & Green, D. H. *J. Petrol.* **12**, 121–147 (1971).
18. Kushiro, I., Syono, Y. & Akimoto, S. *J. geophys. Res.* **73**, 6023–6029 (1968).
19. Green, D. H. *Phys. Earth planet. Inter.* **3**, 221–235 (1970).
20. Frey, F. A., Green, D. H. & Roy, S. D. *J. Petrol.* **19**, 463–513 (1978).
21. Gast, P. W. *Geochim. cosmochim. Acta* **32**, 1057–1086 (1968).
22. Kay, R. W. & Gast, P. W. *J. Geol.* **81**, 653–682 (1973).
23. Schubert, G., Froidevaux, C. & Yuen, D. A. *J. geophys. Res.* **81**, 3525–3540 (1976).
24. Sammis, C. G., Smith, J. C., Schubert, G. & Yuen, D. A. *J. geophys. Res.* **82**, 3747–3761 (1977).
25. Schubert, G & Yuen, D. A. *Tectonophysics* **50**, 197–205 (1978).
26. Yuen, D. A. & Schubert, G. *Geophys. Res. Lett.* **4**, 503–506 (1977).
27. Yuen, D. A. & Schubert, G. *Geophys. J. R. astr. Soc.* **57**, 189–207 (1979).
28. Froidevaux, C. & Souriau, M. *Phil. Trans. R. Soc.* **A288**, 387–392 (1978).
29. Melosh, H. J. *Tectonophysics* **35**, 363–390 (1976).
30. Melosh, H. J. & Ebel, J. *Geophys. J. R. astr. Soc.* **59**, 419–436 (1979).
31. Shaw, H. R. *J. Petrol.* **10**, 510–535 (1969).
32. Shaw, H. R. *Bull. geol. Soc. Am.* **84**, 1505–1526 (1973).
33. Jackson, E. D., Shaw, H. R. & Bargar, K. E. *Earth planet. Sci. Lett.* **26**, 145–155 (1975).
34. Jackson, E. D., Silver, E. A. & Dalrymple, G. B. *Bull. geol. Soc. Am.* **83**, 601–618 (1973).
35. Shankland, T. J. *Oregon Dept Geol. Miner. Ind. Bull.* **96**, 215–218 (1977).
36. Koide, H. & Bhattacharji *Oregon Dept Geol. Miner. Ind. Bull.* **96**, 263–280 (1977).
37. Schwarzer, R. R. & Rogers, J. J. W. *Earth planet. Sci. Lett.* **23**, 286–296 (1974).
38. Miyashiro, A. *Contr. Miner. Petrol.* **66**, 91–104 (1978).
39. Hawkins, J. W. & Natland, J. H. *Earth planet. Sci. Lett.* **24**, 427–439 (1975).
40. DeLong, S. E., Hodges, F. N. & Arculus, R. J. *J. Geol.* **83**, 721–736 (1975).
41. Appleyard, E. C. *Lithos* **7**, 147–169 (1974).
42. Sigurdsson, H., Tomblin, J. F., Brown, G. M., Holland, J. G. & Arculus, R. J. *Earth planet. Sci. Lett.* **18**, 285–295 (1973).
43. Arculus, R. J. *Bull. geol. Soc. Am.* **87**, 612–624 (1976).
44. Arculus, R. J., DeLong, S. E., Kay, R. W., Brooks, G. & Sun, S. S. *J. Geol.* **85**, 177–186 (1977).
45. Price, R. C. & Taylor, S. R. *Contr. Miner. Petrol.* **40**, 195–205 (1973).
46. Smith, I. E. M., Chappell, B. W., Ward, G. K. & Freeman, R. S. *Earth planet. Sci. Lett.* **37**, 230–236 (1977).
47. Johnson, R. W., Wallace, D. A. & Ellis, D. J. *Volcanism in Australia* (Elsevier, Amsterdam, 1976).
48. Johnson, R. W., MacKenzie, D. E. & Smith, I. E. M. *Tectonophysics* **46**, 197–216 (1978).
49. Robin, C & Tournon, J. *Can. J. Earth Sci.* **15**, 1633–1641 (1978).
50. Bevier, M. L., Armstrong, R. L. & Souther, J. G. *Geology* **7**, 389–392 (1979).
51. Heming, R. F. *Lithos* **12**, 173–186 (1979).
52. van Breeman, O., Aftalion, M. & Johnson, M. R. W. *J. geol. Soc. Lond.* **136**, 489–495 (1979).
53. Barth, T. F. W. *Theoretical Petrology* (Wiley, New York, 1962).
54. McBirney, A. R., Aoki, K. *Geol. Soc. Am. Mem. No.* **116**, 523–556 (1968).
55. Baker, B. H., Goles, G. G., Leeman, W. P. & Lindstrom, M. M. *Contri. Miner. Petrol.* **64**, 303–332 (1977).
56. Marsh, J. S. *Earth planet. Sci. Lett.* **18**, 317–323 (1973).
57. Sykes, L. R. *Rev. Geophys. Space Phys.* **16**, 621–688 (1978).
58. Neill, W. M. *Nature phys. Sci.* **245**, 104–107 (1973).
59. Darbyshire, D. P. F. & Fletcher, C. J. N. *Geology* **7**, 545–548 (1979).
60. Opdyke, N & MacDonald, W. D. *Earth planet. Sci. Lett.* **18**, 37–44 (1973).
61. van Breemen, O. & Bowden, P. *Nature phys. Sci.* **242**, 9–11 (1973).
62. Foland, K. A. & Faul, H. *Am. J. Sci.* **77**, 888–904 (1977).
63. Herz, N. *Bull. geol. Soc. Am.* **88**, 101–112 (1977).
64. Amaral, G. *Geochim. cosmochim. Acta* **31**, 117–142 (1967).
65. Dietz, R. S. & Holden, J. C. *J. geophys. Res.* **75**, 4939–4956 (1970).
66. Smith, A. G. *Bull. geol. Soc. Am.* **82**, 2039–2070 (1971).
67. Hsu, K. J. *Nature* **233**, 44–48 (1971).
68. Sorensen, H. (ed.) *The Alkaline Rocks* (Wiley-Interscience, New York, 1974).
69. Vogt, P. R. *J. Geophy. Res.* **86**, 950–960 (1981).
70. Briden, J. C. & Gass, I. G. *Nature* **248**, 650–653 (1974).
71. Gass, I. G., Chapman, D. S., Pollack, H. N. & Thorpe, R. S. *Phil. Trans. R. Soc.* **A288**, 581–597 (1978).
72.. Sengor, A. M. C. *Nature* **279**, 590–593 (1978).
73. Dewey, J. F., Pitman, W. C. III, Ryan, W. B. F. & Bonnin, J. *Bull. Geol. Soc. Am.* **84**, 3137–3180 (1973).
74. Alavi, M. *Geology* **8**, 144–149 (1980).
75. Laubscher, H. & Bernoulli, D. in *The Ocean Basins and Margins* Vol. 4A (eds Nairn, A. E. M., Kanes, W. H. & Stehli, F. G.) 1–28 (Plenum, New York, 1977).
76. Youssef, M. I. *Bull. Am. Ass. Petrol. Geol.* **52**, 601–614 (1968).
77. Swartz, D. H. & Arden, D. D. *Bull. Am. Ass. petrol. Geol.* **44**, 1621–1637 (1960).
78. Burek, P. J. in *Hot Brines and Recent Heavy Metal Deposits in the Red Sea* (eds Degens, E. T. & Ross, D. A.) 59–70 (Springer, New York, 1969).
79. Abdel-Gawad, M. in *Hot Brines and Recent Heavy Metal Deposits in the Red Sea* (eds Degens, E. T. & Ross, D. A.) 25–37 (Springer, New York, 1969).
80. Hancock, P. L. & Kadhi, A. *J. geol. Soc. Lond.* **135**, 339–347 (1978).
81. Brown, G. F. *Phil. Trans. R. Soc.* **A267**, 75–87 (1970).
82. Hsu, K. J. in *The Ocean Basins and Margins* Vol. 4A (eds Nairn, A. E. M., Kanes, W. H. & Stehli, F. G.) 29–75 (Plenum, New York, 1977).
83. Elasser, W. M. *J. geophys. Res.* **76**, 1101–1112 (1971).
84. Pinar-Erdem, N. & Ilhan, E. in *The Ocean Basins and Margins* Vol. 4A (eds Nairn, A. E. M., Kanes, W. H. & Stehli, F. G.) 277–318 (Plenum, New York, 1977).
85. Glennie, K. W. *et al. Bull. Am. Ass. petrol. Geol.* **57**, 5–27 (1973).
86. Turcotte, D. L. & Ahern, J. L. *J. geophys. Res.* **83**, 767–772 (1978).
87. Brown, G. F. *Tectonic Map of the Arabian Peninsula*, Map. AP-2 (Directorate General of Mineral Resources, Jiddah, 1972).

Archaeological evidence for meat-eating by Plio-Pleistocene hominids from Koobi Fora and Olduvai Gorge

Henry T. Bunn

Department of Anthropology, University of California, Berkeley, California 94720, USA

The importance of meat-eating in human evolution has long been a controversial subject[1–4]. The best available evidence of hominid activities between 2 and 1.5 Myr ago is the archaeological record from two East African localities, Olduvai Gorge[5], Tanzania, and Koobi Fora[6], Kenya, which consists of scattered stone artefacts and fragmentary animal bones. The question[7,8] of functional association between juxtaposed artefacts and bones would be largely settled if hominid-induced modifications were present on some bones. Comparative analyses of archaeological bone assemblages from Olduvai Gorge and Koobi Fora and of various modern bone assemblages with known taphonomic histories reveal direct evidence of early hominid butchering and marrow-processing activities.

The surfaces and broken edges of bones from Olduvai and Koobi Fora were examined macroscopically in a strong light. Three general categories of modification were found: (1) hominid-induced damage features, including butchery (cut) marks and hammer-related fracture patterns; (2) carnivore- and rodent-induced damage features, including gnaw marks and tooth-related fracture patterns; (3) weathering and post-depositional alterations, including pitting, grooving and cracking.

A series of very fine linear grooves on bone surfaces constitutes the most unequivocal evidence of hominid involvement with the bones. These grooves are cut marks made by hominids using knife-like stone flakes to remove skin from carcasses, separate articulated bones and detach meat adhering to bones. The grooves occur singly or in multiple sets and vary in length from several millimetres to several centimetres. Typically they

0028-0836/81/250574—04$01.00

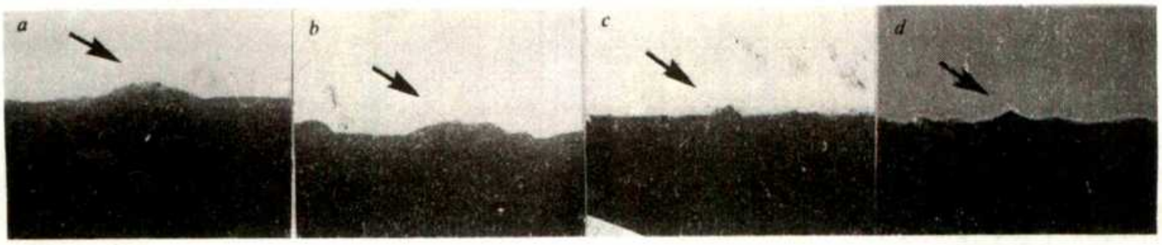


Fig. 1 High-resolution silicone impressions of grooved recent and fossil bone surfaces. Impressions were sectioned transverse to grooves and photographed; ×10.5. Grooves on bone thus appear as positive bumps (arrows) on sectioned impressions of relatively flat bone surfaces. *a*, Broad, rounded cross-sectional shape of recent bone gnawed by carnivore. Width of groove ~1.42 mm. *b*, A series of overlapping small rodent gnaw marks on recent bone, showing flat to slightly rounded cross-sectional shape. Width of individual tooth mark ~0.71 mm. *c*, Much narrower V-shaped cross-section of cut mark made on recent bone using a flint flake. Width of cut mark ~0.36 mm. At FLK Zinj, mean width of 382 cut marks is 0.23 mm, and mean width of 146 carnivore gnaw marks is 0.70 mm. Typical carnivore gnaw mark is thus 3–4 times wider than a cut mark. *d*, Very narrow V-shaped cross-section of cut mark on 1.5-Myr-old fossil hippopotamus femur from site GaJi 5 at Koobi Fora.

are straight-sided and V-shaped in cross-section[9], although some have a slightly flattened bottom, perhaps because of the dullness of the tool being used. These fine cut marks are readily distinguishable macroscopically from carnivore tooth marks, which are typically broader and rounded or U-shaped in section; from rodent gnaw marks, which are typically short, flat or round-bottomed, paired grooves; from irregular, dendritic pitting and grooving that results from the growth of plant roots against bone surfaces; and from faint, shallow linear grooves caused by contact with abrasive sediments (Fig. 1). Moreover, the cut marks have proved indistinguishable macroscopically from undoubted butchery marks, which are, of course, a common feature on archaeological bone debris from more recent periods. On replicas prepared for sectioning and microscopic scrutiny, and for scanning electron microscopy, the marks have also proved indistinguishable from experimentally induced cut marks. Methods for microscopically distinguishing cut marks from other kinds of damage have been developed by Potts and Shipman[10].

Figure 2 shows a particularly well preserved example of this feature from site FxJj50 (ref. 11). The distal end of a very large, extinct bovid humerus has five separate, parallel grooves on its medial surface. Butchering experiments with fresh animal carcasses show that this is a prime location to cut when detaching the humerus from the radio-ulna. The remaining dozen or so examples of butchery marks from FxJj50 occur on limb and rib shafts where the likely aim was to detach meat. At FLK Zinj[5,12], with more complete skeletal representation and better bone preservation, ~300 bones from equid, suid and all bovid genera present, retain cut marks.

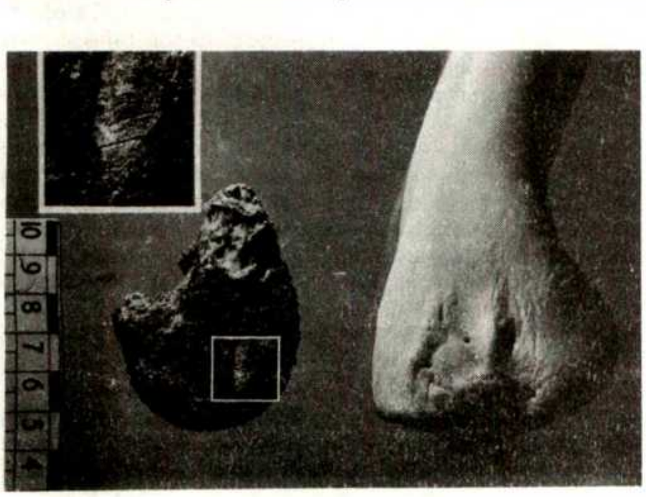


Fig. 2 Medial view of humerus of recent *Connochaetes taurinus* (right) and distal end of humerus of *Megalotragus kattwinkeli* (left) from 1.5-Myr-old site FxJj 50 at Koobi Fora (scale in cm). Inset shows cut marks made by a hominid using a sharp-edged stone artefact; ×1.5. Cut marks are present on 1.6% of FxJj 50 bone fragments and on 8.6% of the FLK Zinj bone fragments that are identifiable to skeletal part.

Similar cut marks have been recognized on fossil bones which did not come from artefact-bearing archaeological sites. These were first recognized on specimens stored in palaeontological collections from Koobi Fora. Because of the high anthropological interest in such evidence, specific searches were made at Koobi Fora, and, in fact, bones with cut marks were recovered during the 1979 season from beds where hominid fossils occur but where artefacts and archaeological sites are sparsely

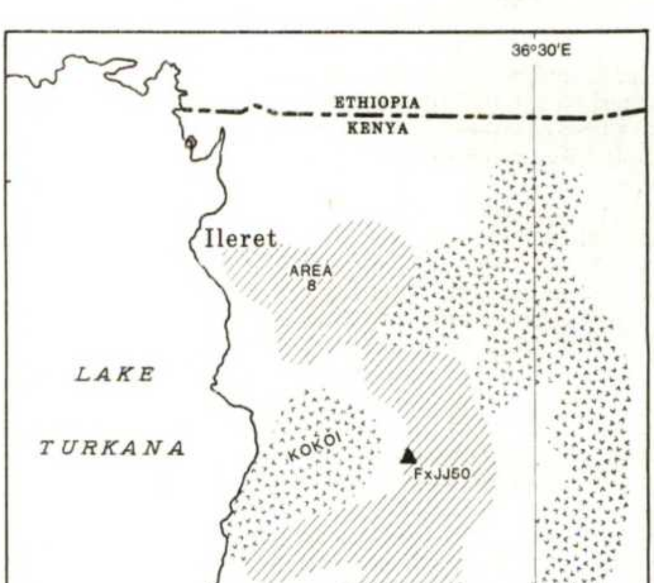


Fig. 3 Map of Koobi Fora study area east of Lake Turkana in northern Kenya, showing selected archaeological sites in relation to lake shore and to outcrops of volcanic and Plio-Pleistocene sedimentary rocks.

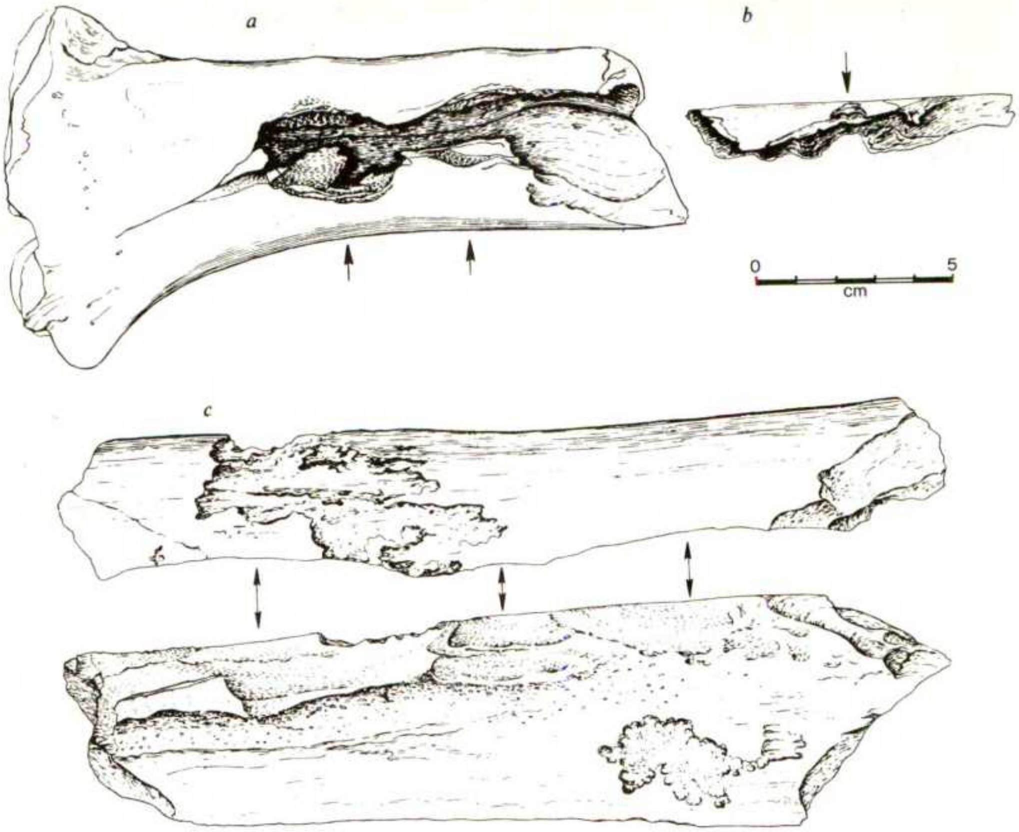


Fig. 4 Recent and fossil bones showing distinctive fracture patterns produced by hammer-stone percussion and carnivore chewing. *a*, Anterior view of proximal end of cow radius broken with hammer-stone. Repeated blows at two adjacent points (arrows) partially split the shaft and produced broad, arcuate indentations and broad bone flakes, several of which adhere to the bone shaft at points of impact. *b*, Recent limb shaft split by hyena chewing. Arrow indicates the point where the tooth contacted and split the shaft, producing relatively small, tooth-sized indentations. *c*, Two views of a very large fossil mammal radius shaft fragment from site FxJj 50: top view shows external surface and broken edge of bone; bottom view shows internal surface of bone with broad flake scars. Arrows mark points of percussion where broad, arcuate indentations indicate probable hammer-stone fracture by hominids (scale in cm). Drawn by B. Isaac.

represented, notably in exposures south of the Koobi Fora base camp (Area 103) and at Ileret (Area 8) (Fig. 3). These beds were deposited near the shoreline of an ancient lake. Source areas of stone raw material are located at least 10–15 km inland, as are almost all previously reported archaeological sites from the Koobi Fora Formation[6]. But at site GaJi5, in Area 103, a surface concentration of ~200 bone fragments but no stone artefacts includes 10 bones, with cut marks, from at least five different animals, including bovid, suid, giraffe and hippopotamus.

Distinguishing between bone assemblages in which a significant amount of fracture has been induced by hominids using hammers and assemblages in which the bones have been broken by chewing, trampling and other processes has proved complex[11,13–19]. Hammer-stone blows produce broad internal flake scars on limb shaft pieces and bone flakes. Damage resulting from breakage by carnivores is similar, but the scars and flakes are much smaller. With carnivore chewing, the edges of cracks along which a bone splits open may again be indented and internally flaked at points where force has been applied, but instead of the broad, arcuate indentation produced by hammer-stone fracture, the indentation approximates the relatively small diameter of the carnivore tooth and thus typically shows a scalloped or denticulate edge with series of tooth-sized indentations along fractured edges (Fig. 4).

Figure 4 shows one example from the excavated assemblage at FxJj50 of a very large mammal radius shaft that was almost certainly broken by hominids using a hammer-stone. At FLK Zinj, the picture is more complex in that over 400 bones show damage attributable to chewing by medium-sized carnivores, and this damage potentially obscures signs of previous hammer-stone breakage by hominids. Although hominids may have broken some of the bones, none of over 50 bones flakes in the assemblage is large enough to preclude derivation from breakage by carnivores. The observations on bone fracture patterns, although less definite, do seem to be consistent with the cut mark evidence in implying significant hominid feeding on animal tissues.

The formation of these early archaeological sites involved a complex interplay of several key factors, including hominid butchering and marrow-processing activities, carnivore scavenging and other animal activities, and various depositional and post-depositional processes. The excavated materials cannot therefore be viewed as a simple, complete record of hominid activities at the sites; rather, the bones and artefacts are a surviving sample of a dynamic system involving processes that added, modified and subtracted items of evidence[11].

The presence of cut marks and percussion marks on significant numbers of bones documents the involvement of early hominids nearly 2 Myr ago in cutting up animal carcasses and breaking open bones, presumably to obtain meat and marrow. In addition to indicating one use of stone artefacts, this direct evidence of early hominid diet allows us to dismiss models of human evolution which do not incorporate meat-eating as a significant component of early hominid behaviour. The documentation of meat-eating and the concentration of bones at particular places by early hominids lends strong support to the food-sharing model proposed by Isaac[20]. The mode of acquisition of this meat, whether hunted, scavenged or both, remains uncertain and requires separate investigation.

These findings also provide a novel technique of mapping the distribution of hominid tool-using and feeding activities across ancient landscapes. In some areas where stone artefacts are too sparse to be detected, fossil bones with cut marks provide evidence of their former presence and usage. Searches for cut marks need to be made in areas that have yielded hominid and other fossil remains from sediments older than 2 Myr, in a time range that may pre-date the habitual discard of quantities of stone artefacts recognizable to archaeologists.

I thank J. D. Clark, G. Isaac, E. Kroll, R. Potts and P. Shipman for comments on an earlier draft of this paper. This research is part of the Koobi Fora Research Project, co-directed by G. Isaac and R. Leakey, of the National Museums of Kenya, and is supported by an NSF grant to G. Isaac.

Received 20 November 1980; accepted 18 March 1981.

1. Dart, R. A. *Transv. Mus. Mem.* **10,** 1–105 (1957).
2. Ardrey, R. *African Genesis* (Collins, London, 1961).
3. Ardrey, R. *The Hunting Hypothesis* (Atheneum, New York, 1976).
4. Zihlman, A. & Tanner, N. in *Female Hierarchies* (eds Tiger, L. & Fowler, H.) (Aldine, Chicago, 1978).

5. Leakey, M. D. *Olduvai Gorge Volume III: Excavations in Beds I and II, 1960–1963* (Cambridge University Press, 1971).
6. Isaac, G. Ll. & Harris, J. W. K. in *Koobi Fora Research Project* Vol. 1 (eds Leakey, M. G. & Leakey, R. E.) (Clarendon, Oxford, 1978).
7. Binford, L. R. *J. anthrop. Res.* **33,** 493–502 (1977).
8. Isaac, G. Ll. *Phil. Trans. R. Soc.* B**292,** 177–188 (1981).
9. Walker, P. L. & Long, J. C. *Am. Antiq.* **42,** 605–616 (1977).
10. Potts, R. & Shipman, P. *Nature* **291,** 577–580 (1981).
11. Bunn, H. *et al. World Archaeology* **12,** 109–136 (1980).
12. Hay, R. L. *Geology of the Olduvai Gorge* (University of California Press, Berkeley, 1976).
13. Sutcliffe, A. J. *Nature* **227,** 1110–1113 (1970).
14. Brain, C. K. in *Human Origins: Louis Leakey and the East African Evidence* (eds Isaac, G. Ll. & McCown, E. R.) (Benjamin, Menlo Park, 1976).
15. Hill, A. P. in *Fossils in the Making: Vertebrate Taphonomy and Paleoecology* (eds Behrensmeyer, A. K. & Hill, A. P.) (University of Chicago Press, 1980).
16. Sadek-Kooros, H. *Am. Antiq.* **37,** 369–382 (1972).
17. Myers, T. P. *et al. Am. Antiq.* **45,** 483–490 (1980).
18. Bonnichsen, R. in *Mammalian Osteo-Archaeology: North America* (by Gilbert, B. M.) (Missouri Archaeological Society, Columbia, 1973).
19. Noe-Nygaard, N. *Paleobiology* **3,** 218–237 (1977).
20. Isaac, G. Ll. *Scient. Am.* **238** (4), 90–108 (1978).

Cutmarks made by stone tools on bones from Olduvai Gorge, Tanzania

Richard Potts* & Pat Shipman†

*Department of Anthropology, Harvard University, Cambridge, Massachusetts 02138, USA
†Department of Earth and Planetary Sciences, The Johns Hopkins University, Baltimore, Maryland 21218, USA

Fossils from Olduvai Gorge, Tanzania, show cutmarks which establish that hominids were using stone tools on animal tissues during the Lower Pleistocene in Africa. We identified cutmarks by elimination of other likely causes of the marks on the bone surfaces, for example, gnawing or chewing by carnivores or rodents, and damage made by tools of excavators or preparators. This was achieved by comparing the marks on the fossils with those produced by known causes on modern bones, using scanning electron microscopy (SEM). Because the fossils occur as part of accumulations of animal remains in relatively undisturbed geological contexts, we conclude that there is a functional association between the stone artefacts and bones at these sites, rather than an accidental, postmortem association[1,2].

The problem in identifying cutmarks produced by stone edges is that similar grooves and marks can be made on bone surfaces by a variety of processes. As part of a study[3] of the effects of various taphonomic processes on bones, the following types of mark were investigated: carnivore tooth scratches, incisal gnawing by rodents or carnivores, root etching, abrasion by windborne and waterborne sedimentary particles, and excavators' and preparators' marks. Each of these alterations to bone surfaces are macroscopically similar to cutmarks. Because the microscopic structure of a bone is a major determinant of its response to potentially destructive forces[4–9], several hundred such marks, produced on modern bones by known events, were inspected by SEM to clarify the interaction between these forces and the microscopic structure of the bone. This study yielded morphological criteria by which cutmarks of various types could be distinguished from the similar marks produced by other events or by normal anatomical features (Fig. 1). Samples are too small at present to develop quantitative criteria.

The controls and a sample of similar marks on fossil bones were inspected on epoxy replicas of bone surfaces; these replicas (which are expected to have a resolution of 0.25 μm) were coated with 200 Å of gold palladium to enhance SEM image quality[10]. Despite the low magnifications used in some cases, the depth of field and image quality of the SEM offered a substantial advantage over a light microscope.

Several types of cutmark were defined for the control sample, on the basis of their mode of production and microscopic characteristics. The term 'cutmarks' is used here as a general term for marks made by stone tools; microscopically, these are elongated grooves with V-shaped cross-sections. Slicing marks are produced by drawing the edge of an artefact across a bone surface in a direction continuous with the long axis of the edge. This creates many fine, parallel striations, visible at a magnification of ×30–×35 or less, within each main groove because the edges of artefacts are not perfectly straight; each small deviation of the edge to one side or the other leaves its own microscopic track. Chopping marks are produced by striking the bone surface with an artefact at a roughly perpendicular angle—these have V-shaped cross-sections and small fragments of bone crushed inwards at the bottom of the main groove. Chopping marks are often broader at the top than slicing marks and do not show fine parallel striations. Scraping marks are formed by drawing an edge across a bone surface in a direction roughly perpendicular to the long axis of the edge. This action results in multiple, fine, parallel striations across a broad area of bone rather than confined to a single, elongated main groove (Fig. 1*a–c*).

No process has yet been discovered which produces marks that mimic slicing, chopping or scraping marks on a microscopic level. Tooth scratches and gnawing of carnivores produce grooves with rounded or flat bases respectively; both lack the fine parallel striations of slicing or scraping marks. Very fine tooth scratches may be as narrow as cutmarks and usually require magnification of ×20–×50 before they can be distinguished from cutmarks. Commonly, rodent gnawing marks are a series of short, very broad, parallel grooves which can be identified with the naked eye. Abrasion by windborne or waterborne sedimentary particles, or by collision of bones with rocks during hydraulic transport, removes external bone surfaces and opens up vascular canals originally lying below the surface, to give a grossly striated appearance. Microscopically, these canals differ from tooth scratches, in showing openings for even smaller vascular canals, and from slicing marks in lacking striations. Scratches produced on fossils with metal tools used for excavation or preparation are not always V-shaped in cross-section and do not shown fine, parallel striations characteristic of cutmarks on fresh bones. The edges of such scratches or grooves are often much more irregular than tooth scratches or gnawing marks (Fig. 1*d–g*).

Numerous fossil bone specimens showing surface marks were found during a study of 12 excavated levels in Beds I and II, Olduvai Gorge[11]. A sample of 75 bone surfaces, some with several marks, was chosen and replicated from fossils at each level. A wide range of skeletal parts, taxa and types of surface damage were included to make the sample as representative as possible. The surface of the fossils was brushed for several minutes with solvents before replication—this was necessary because preservatives completely obscure the microscopic characteristics. The replicas were inspected and photographed at magnifications comparable with those used for the control sample of modern bones.

In the replicated Olduvai sample ($N = 85$ marks on 75 bones), 24% of the specimens possessed marks which were considered to be cutmarks as they more closely resembled those produced by slicing, chopping or scraping than those produced by any other cause so far investigated. Isolated or multiple fine grooves, which could have been confused with cutmarks from a macroscopic inspection, often resembled tooth scratches or narrow gnawing marks much more closely in terms of microscopic detail. Table 1 gives the numbers of fossils with each type of mark; Fig. 2*a* shows a series of slicing and scraping marks on a fossil.

Significantly, several bones had both slicing marks and tooth scratches, including three specimens which showed intersections between these types of mark (Fig. 2*b*, *c*). Thus it is clear that, at least in some cases, hominids and carnivores were able to use the same parts of carcasses in turn. Furthermore, slicing marks are present on at least one specimen from FLK NN, Level 2, although no stone tools are known from that site. One possible interpretation of this evidence is that carnivores scavenged this bone from a hominid campsite and removed it to another area.

0028-0836/81/250577—04$01.00

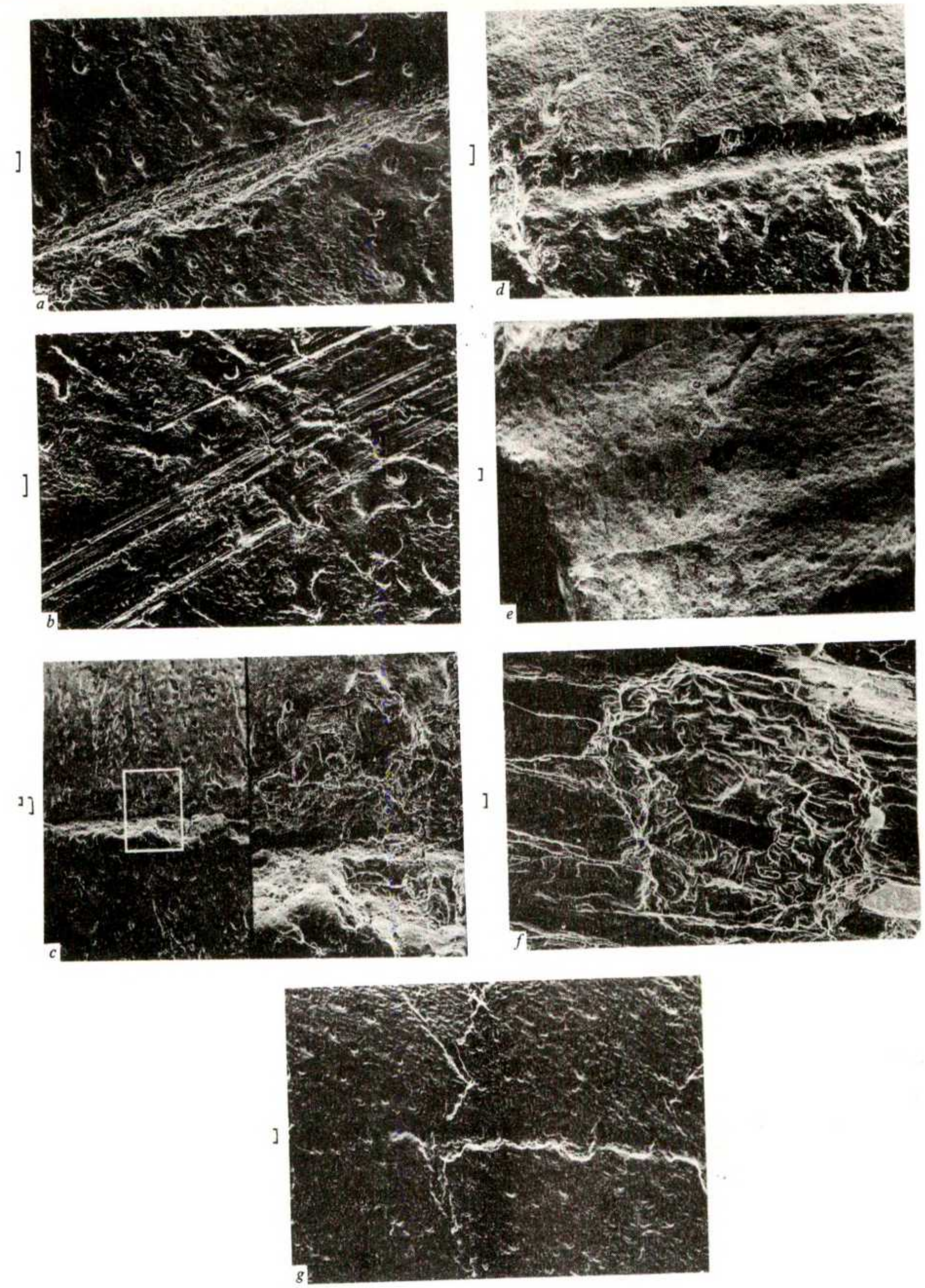


Fig. 1 Micrographs of various tooth and tool marks. The bar or bars to the left of each figure represent 100 μm. *a*, A slicing mark on a modern bone, showing the fine parallel striations within the main groove. *b*, A scraping mark on a modern bone, showing fine striations across a broad area. *c*, A chopping mark on a modern bone, shown at two magnifications. Left, at low power; note the broad, V-shaped groove lacking fine striations. Right, at higher power; note the fragments of bone crushed inwards at the bottom of the groove. *d*, A carnivore tooth scratch on a modern bone. Although this is the functional equivalent of a slicing mark, there are no fine striations. *e*, A rodent gnawing mark on a modern bone. This is the functional equivalent of a scraping mark, but gnawing produces only broad, shallow grooves without fine striations. *f*, A puncture produced by a carnivore canine on a modern bone. Punctures are functionally equivalent to chopping marks and are similar in that fragments of bone are crushed inwards. However, punctures are typically rounded in outline, unlike chopping marks. *g*, A preparator's marks on a fossil bone, produced by an action similar to slicing. The lack of striations distinguish it from a slicing mark and the irregular edges of the groove distinguish it from a tooth scratch.

Table 1 Data on frequencies and locations of different types of ma[illegible]n fossils from Beds I and II, Olduvai Gorge

Type of mark	No. of specimen	Locations	Taxa	Sites
Slicing	13	Metatarsal shaft (2), radius shaft (2), limb shaft fragment (2), humerus shaft, metacarpal shaft, proximal ulna, distal metacarpal, rib shaft, tibial crest, proximal radius	*Kobus* [illegible]*oidalis*, *Equus oldowa*[illegible]*is*, giraffid indeterminate, small b[illegible] indeterminate, large bovid i[illegible]erminate, bovid indeter[illegible]te, medium/large mammal indeterm[illegible]te	DK Levels 2 and 3, FLK NN Levels 2 and 3, FLK 'Zinj', MNK Main
Scraping	5	Scapula blade, humerus shaft, radius shaft, metatarsal shaft, limb shaft fragment	*Kobus sig*[illegible]*oidalis*, *Equus oldowaye*[illegible]*s*, giraffid indeterminate, medium m[illegible]mmal indeterminate, large mammal indeterminate	DK Level 3, FLK 'Zinj', MNK Main
Chopping	5	Metacarpal shaft (2), metapodial shaft, proximal phalanx shaft, limb shaft fragment	*Parmularius altidens*, *Equus oldowayensis*, bovid indeterminate, mammal indeterminate	DK Levels 2 and 3, FLK NN Level 2, MNK Main
Tooth/gnawing	48	Limb shaft fragment (12), proximal rib (6), humerus shaft (4), rib shaft (3), femur shaft (3), tibial crest (2), metacarpal shaft (2), proximal metacarpal (2), scapula glenoid, proximal ulna, radius shaft, radio-ulna shaft, mandibular corpus, hyoid, parietal vault, proximal femur, acetabulum, distal tibia, fibula shaft, metatarsal shaft, cuneiform, indeterminate fragment	*Kobus sigmoidalis*, *Megalotragus kattwinkeli*, *Elephas recki*, *Deinotherium* sp.; *Equus oldowayensis*, *Cercocebus*, *Homo habilis*, Tragelaphini indeterminate, Reduncini intermediate, Antilopini indeterminate, medium/large bovid indeterminate, medium/large mammal indeterminate, large mammal indeterminate, very large mammal indeterminate, mammal indeterminate	DK Levels 2 and 3, FLK NN Level 2 and 3, FLK 'Zinj', FLK N Level 6, FLK N Deino. Level, MNK Main, TK Upper Floor
Preparation/ excavation	7	Scapula blade, cervical vertebra, rib shaft, astragalus, cuneiform limb shaft fragment, indeterminate fragment	*Elephas recki*, Alcelaphini indeterminate, medium bovid indeterminate, medium/large Carnivora indeterminate, mammal indeterminate	FLK 'Zinj', FLK NN Level 6
Undetermined	7	Rib shaft, distal tibia, patella, metatarsal shaft, lateral cuneiform, limb shaft fragment, indeterminate fragment	*Equus oldowayensis*, small bovid indeterminate, very small mammal indeterminate, large mammal indeterminate	DK Level 3, FLK NN Level 2, FLK N Level 6, MNK Main

Seventy-five surfaces were examined; 10 of these possessed marks which classified into two different types. Numbers in parentheses after skeletal location signify the number of times different specimens showed a particular type of mark in that location.

Another likely explanation is that hominids were the primary agent responsible for accumulating and damaging the bones at FLK NN, Level 2, but did not discard stone tools there.

Cutmarks were commonly found on fragments of fossilized bone, particularly on pieces of limb bone shaft, which suggests that the marks occurred during defleshing or dismembering of carcasses with stone tools (Table 1). However, the cutmarks on bones identifiable to skeletal element and species are neither frequent enough nor do they have sufficient regularity of location to suggest a consistent butchery pattern such as can be reconstructed for various species' carcasses much later in time in the New World[12,13]. To determine the number of cutmarks in a single taxon, all marks found on all postcrania ($N = 138$) of one species, *Equus oldowayensis*, were examined; only 2.9% showed cutmarks. If one or more standard strategies for butchery of equids with stone tools existed, that sample is too small to reveal them.

Specimens from identified skeletal elements were divided into meat-bearing and non-meat-bearing bones. Table 2 shows a preferential occurrence of tooth marks on meat-bearing bones; cutmarks occur about equally on both kinds of bone. Evidently, hominid use of stone tools on bones was not directed solely towards meat acquisition. Carnivores and rodents, however, focused more directly on meat or on tissues inside meat-bearing bones.

Table 2 Occurrence of stone tool cutmarks and animal tooth marks on meat-bearing bones (major limb bones and axial elements) as opposed to non-meat-bearing bones (metapodials, podials, phalanges)

	Stone cutmark	Tooth mark
Meat-bearing bone	10	27
Non-meat-bearing bone	9	7

The sample consists of 53 fossil surfaces from Olduvai, identified to skeletal element. $\chi^2 = 4.15$, degrees of freedom = 1, $0.025 < P < 0.05$.

Our evidence supports Leakey's conclusion[1] that there is a direct causal association between the stone artefacts and fossilized bones at many Olduvai sites. This evidence is important because it comes from some of the oldest known excavated archaeological levels: for example, DK I, Level 3, is older than 1.79 ± 0.03 Myr (ref. 2). Furthermore, the cutmarks indicate that hominids processed carcassses to obtain animal material such as meat, ligament or bone. However, this evidence does not bear directly on questions about the frequency of meat-eating and hunting among early hominids[14–17]. Moreover, the occurrence of both cutmarks and tooth marks on the bones documents an overprinting of taphonomic events; both hominids and other animals modified the bones in these assemblages. Therefore we cannot attribute the observed patterns of faunal and skeletal representation solely to hominid activity.

This study establishes: (1) a technique for distinguishing cutmarks made with stone artefacts from similar marks not involving early hominid activities; (2) that cutmarks were produced by hominids using stone artefacts on animal tissues at these sites at least as early as 1.79 Myr ago; (3) that these assemblages were exposed to multiple agents of damage and accumulation and thus their composition cannot be taken as indicative solely of hominid behaviours; (4) that not all bones altered by hominids will be found in association with stone tools; and (5) that hominids and carnivores were in competition for carcasses or bones, perhaps to obtain different substances.

We thank the governments of Kenya and Tanzania, Mary and Richard Leakey, and the National Museums of Kenya for permission to study fossils and bones in their care; and Alan Walker and Glynn Isaac for valuable comments. We also thank

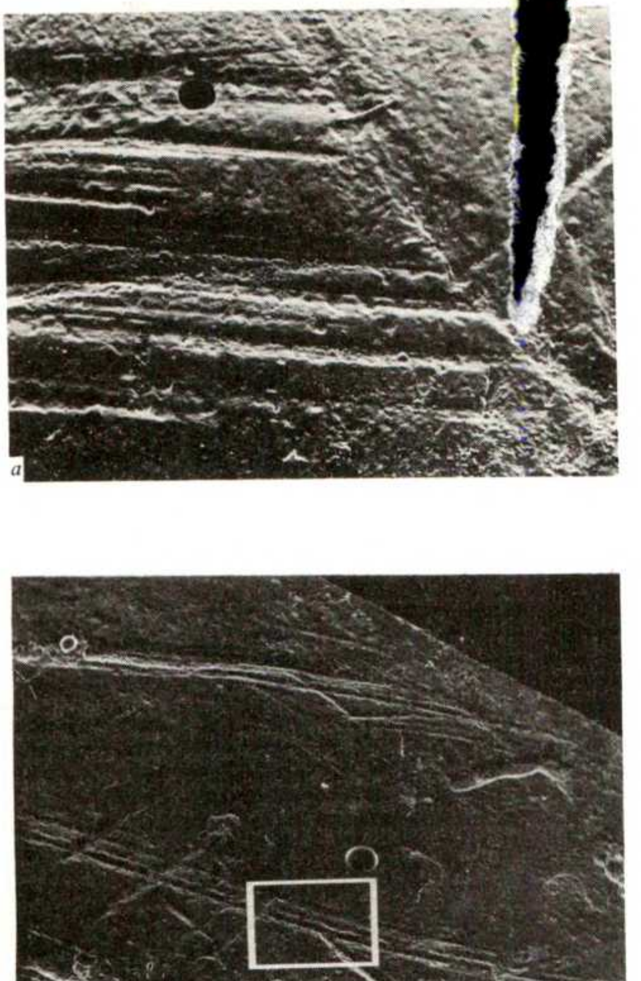


Fig. 2 Marks on Olduvai fossils. *a*, A series of slicing and scraping marks run roughly horizontally across the micrograph. A slicing mark made later in time runs diagonally downwards across the earlier marks. The specimen is an indeterminate shaft fragment from DK I, Level 3. *b*, Two slicing marks run roughly horizontally across this micrograph. The lower one is overlaid by several tooth scratches which must have been made after the slicing marks. The box encloses the intersection of two of these tooth scratches with the slicing mark. The specimen is an equid tibia from FLK Zinj. *c*, A tracing of the micrograph in *b*, with various features labelled.

Linda Perez, Wendy Bosler and Karen Lee Davis for technical assistance. This work was funded by NSF grant BNS 78 19174 to R.P., the Boise Fund (to R.P. and P.S.) and a NIH biomedical research grant 5 S07RR07041-13 to P.S.

Received 18 December 1980; accepted 9 April 1981.

1. Leakey, M. D. *Olduvai Gorge* Vol. 3 (Cambridge University Press, London and New York, 1971).
2. Hay, R. L. *Geology of the Olduvai Gorge* (University of California Press, Berkeley, Los Angeles and London, 1976).
3. Shipman, P. *Ann. N.Y. Acad. Sci.* (in the press).
4. Lanyon, L. E. & Smith, R. N. *Acta orthop. scand.* **41,** 238–248 (1970).
5. Evans, F. G. *Mechanical Properties of Bone* (Thomas, Springfield, 1973).
6. Lakes, R. & Saha, S. *Science* **208,** 501–503 (1979).
7. Bonnichsen, R. *Pleistocene Bone Technology in the Beringium Refugium* Archaeological Surv. Canada **89** (1978).
8. Sadek-Kooros, H. *Am. Antiq.* **37,** 369–382 (1972).
9. Brain, C. K. in *Human Origins* (eds Isaac, G. Ll. & McCown, E.) 97–116 (Benjamin, Menlo Park, 1976).
10. Walker, A. *Phil. Trans. R. Soc.* B**292,** 56–64 (1981).
11. Potts, R. (in preparation).
12. Guilday, J. E., Parmalee, P. W. & Tanner, D. P. *Penna Arch.* **32,** 59 (1962).
13. Frison, G. C. *Plains Anthrop.* **16,** 258 (1971); *The Casper Site* (Academic, New York, 1974).
14. Isaac, G. Ll. *Scient. Am.* **238** (4), 90–108 (1978).
15. Tanner, N. & Zihlman, A. *Signs* **3,** 585–608 (1976).
16. Binford, L. R. *Nunamuit Ethnoarchaeology* (Academic, New York, 1978).
17. Jolly, C. J. *Man* **5,** 5–20 (1970).

Chondroitin sulphate from fossilized antlers

J. E. Scott & E. W. Hughes

Department of Medical Biochemistry, University of Manchester
Medical School, Oxford Road, Manchester M13 9PT, UK

If biopolymers could be isolated from archaeological specimens there would be good prospects of correlating the gene products from different populations of the same species, for example, according to their immunological specificities. Soft tissues are very poorly preserved after only decades, but the macroscopic appearance of bone and antler can remain unchanged for thousands of years. Methods which allow the extraction of relatively undegraded biopolymers from such materials would extend by several orders of magnitude the period from which biopolymers could be identified[1]. We report here that chondroitin sulphate, a polysaccharide characteristic of connective tissue and with chemical similarities to immunologically active glycosaminoglycans, has been isolated in good yield from fossilized antlers (3,000–130,000 yr old) which were demineralized by a new procedure. There are now prospects for using biochemical and immunological methods *in situ* and on isolated biopolymers to investigate the composition and mobility of populations of species in the past.

It is first necessary to remove the mineral. Previous methods have used acids or aqueous media containing EDTA but neither can be recommended to secure high yields of undegraded biopolymers of the types which offer maximum prospects of serological specificity, that is, glycoproteins. Aqueous EDTA demineralizes slowly, allowing degradation of water-soluble glycosaminoglycans extracted into solution. We developed a method[2], using EDTA in organic solvents, which leaves most water-soluble polymers *in situ*. Such media are much less reactive than aqueous environments. Several glycosaminoglycan components of bone and antler were identified[3] which had escaped detection in aqueous EDTA-treated samples. Structural proteins and glycoproteins were demonstrable by immunofluorescent techniques in bone and antler demineralized by the new method (J. E. S., R. Timpl and K. von der Mark, unpublished results). It seemed that the application of this method to fossils might give a maximum yield of minimally degraded material.

Table 1 Analyses per g of original antler

Sample	Age (yr)	Ca (nmol)	Hydroxyproline (mg)	Uronic acid in isolated polysaccharide (μg)	Hexosamine in isolated polysaccharide (μg)
Modern		5.4*	42	300	
a	3,000–5,000	8.4*	1.65	28.9	38
b	3,800		1.3	21	18.7
c	38,850†	6.7*	17.4	43.4	45‡
d	12,000	6.1§	14.8	110	85
e	130,000	8.2§	0.77	7.8	4.4

* EDTA titration.

† Radiocarbon date (supplied by Mr R. Burleigh).

‡ Galactosamine, by ion exchange chromatography (from Dr A. Shuttleworth).

§ Atomic absorption (supplied by Dr G. A. Lumb).

Two of the samples of antler used were; *a*, red deer (*Cervus elaphus*) from Grimes Graves, estimated as 3,000–5,000 yr old, and *c*, reindeer (*Rangifer tarandus*) from Creswell Crags, Derbyshire, radiocarbon-dated as 38,850 ± 2,500 yr old. The probable age of sample *b* (red deer antler, find no. 806 from Grimes Graves, Greenwells Pit, in contact with chalk rubble) was 3,800 yr (ref. 4); that of samples *d* (*Megaceros giganteus*, from Antrim bog) and *e* (*Cervus elaphus*) were ~12,000 and 130,000 yr, respectively. These samples were decalcified for 4–8 weeks in 80–20% ethanol–water solutions containing 0.5 M trimethylammonium EDTA at *p*H 8 and 37 °C (ref. 1). The decalcified material was digested with papain at 65 °C and acidic polysaccharides were isolated from the digest using cetylpyridinium as described elsewhere[5], and electrophoresed[3]. Hydroxyproline, hexosamine and uronic acid determinations, and chondroitinase digestion of the isolated polysaccharides were done using methods described elsewhere[6–9]. Table 1 lists the results.

Electrophoresis, chondroitinase digestion, hexosamine and uronic acid determinations indicated a considerable yield of very good quality chondroitin sulphate, which agrees well with the finding of chondroitin-4-sulphate in contemporary samples of antler[3]. The yield from antler *c* was ~15% of that from modern antler, on a weight basis.

The success of the methods applied to the fossilized antlers suggests that much older material might yield significant quantities of glycosaminoglycan-type material. Although chondroitin sulphate represents side chains rather than the whole proteoglycan, it would be remarkable if every vestige of the polypeptide had been detached in the fossil. Indeed, amino acids (including serine) typical of the polypeptide core around the attachment region, which are normally found in papain-digested chondroitin proteoglycan, were present in the hydrolysate of chondroitin sulphate from sample *c*. It is possible that some of the chondroitin sulphate constituted part of macromolecules which could be defined as proteoglycan.

The conditions in which the sample was lying should influence the yield of glycosaminoglycan. Sample *c*, from a dry and slightly alkaline limestone cave, was well preserved. More interesting is the excellent yield from sample *d*, which was considerably less mineralized than the others. It is possible that components of the bog water had had a preservative or fixative action.

We observed the histological distribution of chondroitin sulphate in decalcified sample *c* which had been embedded in wax and cut on conventional microtomes, using standard techniques (Alcian Blue–critical electrolyte concentration (CEC) staining according to the method of Scott and Dorling[10]). It should be possible to apply immunofluorescence methods to decalcified fossils, as we have to modern antler (J. E. S., R. Timpl and K. von der Mark, unpublished results). Other calcified tissues should be expected to respond, thus making the 'molecular museums' of skeletal tissues available to modern immunological procedures.

Antibodies have been applied to fossils to identify proteins (particularly collagen and albumin[11]) and it has proved possible to distinguish between fossil species. Glycoproteins carry antigenic reactivities of very high specificity (for example, blood groups) which reside in the glycosaminoglycan portion. The demonstration that a glycosaminoglycuronan structure survives in good yield for tens of thousands of years suggests that it will be rewarding to look for similar compounds, either by immunological or newly developed chemical methods of the required specificity and sensitivity (see part 2 of ref. 2, 6–201). Intraspecies information on populations from the distant past should then be available.

We thank Mr Burleigh for sample *b* and for radiocarbon dating, Professor Martin Jope for samples *d* and *e*, and Dr Prag for samples *a* and *c*; all provided valuable discussion. We also thank Dr A. Shuttleworth for the amino acid determination.

Received 29 September 1980; accepted 8 April 1981.

1. Jope, E. M. *Biogeochemistry of Aminoacids* (ed. Hare, P. E.) 23 (Wiley, New York, 1980).
2. Scott, J. E. & Kyffin, T. W. *Biochem. J.* **169,** 697–701 (1978).
3. Scott, J. E., Kyffin, T. W. & Hughes, E. *Proc. 5th. Int. Symp. Glycoconjugates* (eds Schauer, R. *et al.*) 566 (Georg Thieme, Stuttgart, 1979).
4. Burleigh, R., Hewson, A., Meeks, N., Sieveking, G. & Longworth, I. *Radiocarbon* **21,** 41–47 (1979).
5. Scott, J. E. *Meth. biochem. Analysis* **8,** 145–197 (1960).
6. Woessner, J. F. Jr *Archs Biochem. Biophys.* **93,** 440–447 (1961).
7. Bitter, T. & Muir, H. M. *Analyt. Biochem.* **4,** 330–334 (1962).
8. Kraan, J. G. & Muir, H. M. *Biochem. J.* **66,** 55p (1957).
9. Saito, H., Yamagata, T. & Suzuki, S. *J. biol. Chem.* **243,** 1536–1542 (1968).
10. Scott, J. E. & Dorling, J. *Histochemie* **5,** 221–233 (1965).
11. Lowenstein, J. M. *Phil. Trans. R. Soc.* (in the press).

Monoclonal antibodies that distinguish between New World species of *Leishmania*

Diane McMahon Pratt & John R. David

Department of Medicine, Harvard Medical School, and Department of Rheumatology and Immunology, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA

The New World *Leishmania* species[1] *Leishmania braziliensis*, which causes mucocutaneous leishmaniasis, can be distinguished from *Leishmania mexicana*, which causes cutaneous leishmaniasis, by the buoyant density of nuclear and kinetoplast DNA[2], electrophoretic mobility of different isoenzymes[3–6], by differing characteristic growth patterns in the sandfly and hamster and *in vitro*[7,8], and by size in electron microscopy[9,10]. However, there is no available method for rapid and accurate diagnosis of fresh isolates of New World leishmaniasis species. Furthermore, immunological cross-reaction of *Leishmania* with *Trypanosoma cruzi*, which is co-endemic with the former, has posed a serious problem in the use of serodiagnostic tests for leishmaniasis[11–13]. The ability to distinguish at primary isolation between the species causing cutaneous and mucocutaneous leishmaniasis would make it possible to concentrate on treating those patients likely to suffer serious disfiguring disease. To develop such a specific diagnostic test, we have produced monoclonal antibodies[14] specific for either the *L. mexicana* or *L. braziliensis* species. One of these monoclonal antibodies is specific for a subgroup of *L. braziliensis*; none is cross-reactive with *T. cruzi* (Y strain) epimastigotes. These monoclonal antibodies should be useful in the taxonomic identification of different species of New World leishmaniae as well as for the direct diagnosis of leishmaniasis.

Female BALB/c mice were injected with membrane-rich preparations isolated from either *L. braziliensis panamanensis* or *L. mexicana amazonensis* promastigotes. The membrane preparations were obtained by disrupting the promastigotes by

0028-0836/81/250581—03$01.00

nitrogen cavitation, followed by subfractionation by differential centrifugation and purification on sucrose density gradients. The method of fusion was that of Kennett *et al.*[15]. Antibody production was assessed between days 14 and 17 after fusion by an indirect radioactive binding assay using leishmanial membranes hydrophobically fixed to PVC microtitre plates. Cultures secreting antibody were then assessed for cross-reactivity with *Leishmania hertigi* or *Leishmania enrietti* antigens. Selected wells were then cloned by limiting dilution.

The average values for six fusions were: 73% of the total wells plated showed hybrid growth; 18% of these hybrid-containing wells produced antibodies to *Leishmania* antigens and 0.6% of the hybrid-containing wells were producing species-specific antibody. The specificity of the monoclonal *Leishmania* antibodies produced is shown in Fig. 1. Monoclonal antibodies IV-3B9, V-1A12 and IX-1G3 are cross-reactive positive controls. The amount of radioiodinated rabbit anti-mouse immunoglobulin that bound to plates containing the *L. braziliensis panamanensis* membrane antigens, previously incubated with the *L. braziliensis* species-specific hybrid supernatants, was 17 to 21 times greater than that which bound after preincubation with control medium. After incubation with *L. mexicana* species-specific hybrid supernatants, the amount of radioiodinated rabbit anti-mouse immunoglobulin that bound to plates containing *L. mexicana amazonensis* membrane antigens was 6.1 to 38 times greater than that which bound to the medium controls. When *L. hertigi* membrane antigens were used, the amount of radioiodinated rabbit anti-mouse immunoglobulin bound in the case of either *L. braziliensis* or *L. mexicana* species-specific supernatant was only 0.7 to 3.0 (average value 1.3) times that bound in the medium controls. Similar specificity was seen if the *L. braziliensis* hybrids were assayed for reactivity with *L. enrietti*, *L. mexicana amazonensis* or *L. hertigi*, or if the *L. mexicana* hybrids were assayed for reactivity with *L. braziliensis braziliensis* or *L. hertigi* (data not shown).

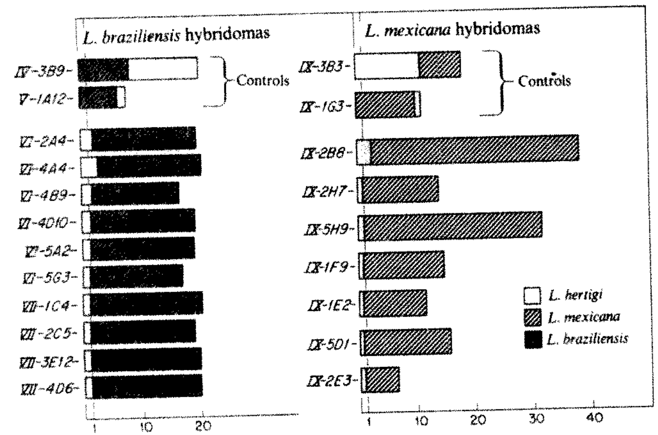


Fig. 1 An indirect radioimmune binding assay was used to determine the reactivities of the culture supernatants from *L. braziliensis*-specific hybrids with *L. braziliensis panamanensis* and *L. hertigi* membrane antigens and the reactivities of culture supernatants from *L. mexicana*-specific hybrids with *L. mexicana amazonensis* and *L. hertigi* membrane antigens. Antibodies produced by the hybridomas were measured by radioactive binding assay. Briefly, the screening assay was as follows. Sonicated leishmanial membranes diluted in phosphate-buffered saline (PBS) containing 0.02% NaN_3 were plated in Cooke 'U'-bottom 96-well PVC plates overnight at 5 °C. The plates were washed six times in PBS containing 5% horse serum at 0.02% NaN_3. Culture medium supernatants containing secreted antibodies were incubated with the antigen plates for a minimum of 2 h at 5 °C. Immune mouse serum (diluted 1 : 20) and cell culture medium were used as positive and negative controls, respectively. The plates were again washed six times with PBS containing 5% horse serum and 0.02% NaN_3. Affinity-purified ^{125}I-labelled rabbit $F(ab')_2$ anti-mouse immunoglobulin (5–10 ng per well; 10^5 c.p.m.) was added to the wells and incubated for 1 h at 0 °C. Excess antibody was again removed by washing five times. The plates were then air dried and the radioactivity bound to each well measured using a Packard Auto-Gamma counter. The background (c.p.m. bound) was 510 for *L. braziliensis panamanensis* and 217 for *L. hertigi*, in the case of the *L. braziliensis* hybridomas. The background (c.p.m. bound) was 340 for *L. mexicana amazonensis* and 440 *L. hertigi* for the *L. mexicana* hybridomas.

Table 1 Relative reactivities of *Leishmania* monoclonal antibodies with epimastigotes of *T. cruzi*

		Ratio of bound antibody to bound control		
Code	Clone	*L. mexicana amazonensis*	*L. braziliensis panamanensis*	*T. cruzi*
M-1	IX-2B8-D2	21.0	—	1.1
M-2	IX-2H7-E10	23.0	—	0.7
M-3	IX-5H9-C1	17.0	—	0.9
M-4	IX-1F9-A9	10.6	—	0.9
M-5	IX-1E2-D12	8.7	—	1.3
M-6	IX-2E3-C10	2.4	—	1.0
M-7	IX-5D1-C10	1.4	—	1.1
B-1	VI-4A4-A8	1.0	40.0	1.4
B-2	VI-4B9-D10	1.2	59.7	1.0
B-3	VI-4D10-D12	1.5	39.9	1.3
B-4	VI-2A5-A4	1.4	42.0	1.2
B-5	VII-2C5-C5	1.2	49.0	0.8
B-6	VII-1C4-H6	1.4	45.0	0.9
B-7	VI-2A4-E3	1.2	40.7	1.7
B-8	VII-3E12-D3	1.2	43.0	0.9
B-9	VII-4F11-A6	1.6	38.8	1.0
B-10	VII-4D6-E8	1.2	62.9	0.8
B-11	VI-5G3-F3	1.1	50.0	0.9
	V-1A12-G2	21.5	14.8	14.2
	V-3G5-C6	21.5	8.2	12.4

Reactivities of the culture supernatants of cloned hybrids with external membrane components of *Leishmania* were determined in an indirect radioactive binding assay using glutaraldyhyde-fixed *Leishmania*. Glutaraldehyde fixation of the various *Leishmania* strains was carried out as described elsewhere[16]. The binding assay procedure was essentially as described in Fig. 1 legend except that 'V'-bottom microtitre plates (Cooke) were used and the *Leishmania* (3–5×10^6 per well) were washed only three times after being pelleted at 400*g* for 10 min after each incubation with antibody. On completion of the assay, the *Leishmania* were resuspended after the final wash in 100 μl of 0.5 M NaOH and 60 μl was taken for radioactivity measurement.

The reactivities of the culture supernatants of cloned hybrids with external membrane components of *Leishmania* were determined in an indirect binding assay using glutaraldehyde-fixed *Leishmania* (Table 1). All the specific hybridoma antibodies tested, except M-7, bound to the appropriate glutaraldehyde-fixed *Leishmania*. As almost all the antibodies of interest appeared to react with external surface components, glutaraldehyde-fixed organisms were used in the indirect binding assay. In addition, this assay was chosen because it avoids proteolysis, which can occur during membrane isolation. In the case of clone M-7, the antigen recognized is either destroyed by glutaraldehyde or the antigen is an internal membrane-associated component.

No significant reactivity was found for any of the species-specific monoclonal antibodies with *T. cruzi* (strain Y) epimastigotes (Table 1). Clones V-1A12-G2 and V-3G5-C6 produce cross-reactive *Leishmania* monoclonal antibodies that are clearly cross-reactive with the *T. cruzi* epimastigotes. The levels of reactivity of these clones with the *T. cruzi* epimastigotes are comparable to those with *L. mexicana amazonensis* and *L. braziliensis panamanensis*.

The species-specific monoclonal antibodies were assessed for their reactivity against a panel of *L. mexicana* and *L. braziliensis* strains (Table 2). Each strain designation had been recently determined by isoenzyme analysis. Isolates designated as WR303, Tres Bracos 04 and Tres Bracos 11 are *L. mexicana*; M2903 and Tres Bracos 05, 08 and 13 are *L. braziliensis*.

Table 2 Comparative reactivities of monoclonal antibodies to strains of *L. mexicana* and *L. braziliensis*

		Ratio of bound antibody to bound control						
		L. mexicana			*L. braziliensis*			
Code	Clone	WR303	Tres Bracos 04	Tres Bracos 11	Lbb M2903	Tres Bracos 05	Tres Bracos 13	Tres Bracos 08
M-1	IX-2B8-D2	19.6	14.7	21.1	0.8	1.0	1.1	0.9
M-2	IX-2H7-E10	21.4	14.8	17.9	0.6	0.9	0.7	0.7
M-3	IX-5H9-C1	16.0	7.2	10.3	0.6	0.9	0.8	0.7
M-4	IX-1F9-A9	9.3	5.5	7.7	0.6	0.9	0.9	0.8
M-5	IX-1E2-D12	8.1	5.1	7.3	0.8	1.2	1.2	1.0
B-1	VI-4A4-A8	1.0	2.8	ND	15.5	ND	ND	ND
B-2	VI-4B9-D10	1.2	0.9	1.2	7.3	29.5	14.0	13.1
B-3	VI-4D10-D12	1.5	1.9	1.5	60.4	30.9	13.1	12.8
B-4	VI-2A5-A4	1.4	1.2	2.6	7.5	20.4	10.7	6.6
B-5	VII-2C5-C5	1.1	0.7	1.3	78.6	55.2	7.4	14.5
B-6	VII-1C4-H6	1.4	1.1	2.1	5.9	30.3	7.6	10.9
B-7	VI-2A4-E3	1.2	1.6	12.8	10.8	30.6	16.4	14.4
B-8	VII-3E12-D3	1.1	0.9	9.6	63.8	27.5	9.7	11.8
B-9	VII-4F11-A6	1.6	1.2	13.7	4.3	24.5	10.0	93
B-10	VII-4D6-E8	1.2	0.9	5.2	80.2	29.6	14.4	15.7
B-11	VI-5G3-F3	1.1	0.7	1.7	0.9	1.4	5.4	0.4

The reactivities of the culture supernatants of cloned hybrids with glutaraldehyde-fixed isolates of *L. mexicana* and *L. braziliensis* were determined in an indirect radioactive binding assay as described in Table 1 legend. Walter Reed strain WR303 is *L. mexicana amazonensis* isolate LV72 from Liverpool; Tres Bracos strains 04, 05, 08, 11 and 13 are Universidade de Brasilia stocks LTB-0016 (Irene), LTB0042 (Jurandir), LTB0014 (Mucosa), LTB0055 (Jose) and LTB0048 (Elaine) respectively. Strain Lbb M2903 is the reference *L. braziliensis braziliensis* isolate of Dr R. Lainson. Isoenzyme analyses of each strain were performed by Dr M. Miles. ND, not determined. Results of note are demarcated by boxes.

Clones M-1 to M-5 are reactive with the *L. mexicana* strains but unreactive with the *L. braziliensis* strains. In contrast, clones B-1 to B-6 and B-11 are reactive with the *L. braziliensis* strains, but not *L. mexicana* strains. Clones B-7 to B-10 are reactive with all the *L. braziliensis* strains but not the *L. mexicana* strains *L. mexicana amazonensis* WR303 and Tres Bracos 04; however, these are reactive with the *L. mexicana* strain Tres Bracos 11. Clone B-11 seems to be specific for a subgroup of *L. braziliensis*, reacting only with the strains *L. braziliensis* (Tres Bracos 13) and *L. braziliensis panamanensis.*

The cross-reactivity of clones B-8, B-9 and B-10 may be useful in studying the subspeciation of *L. mexicana* strains. It is interesting that the *Leishmania* strains Tres Bracos 04 and Tres Bracos 11 were isolated from the same area of Brazil (Tres Bracos). In addition, clone VI-5G3-F3, reacting with *L. braziliensis* Tres Bracos 13, may be useful in studies of subspeciation of *L. braziliensis.* Work is being done to develop these monoclonal antibodies as clinical reagents for direct diagnosis of leishmaniasis.

We thank Elizabeth Bennett for technical assistance and Dr Larry Hendricks for advice and for providing the *L. mexicana amazonensis* (WR303) and *L. braziliensis panamanensis* (WR120) strains. We also thank Dr R. Lainson for the *L. braziliensis braziliensis* reference strain M2903, Dr Michael Miles for isoenzyme analysis, Dr Rodney Hoff for *T. cruzi* epimastigotes and Dr Philip Marsden for the Tres Bracos strains 04, 05, 08, 11 and 13 and for discussions which directly led to the work. This work was supported by NIH grant AI-16479 and by grants from the Rockefeller foundation and CNPQ, the National Research Council of Brazil.

Received 22 December 1980; accepted 25 March 1981.

1. Marsden, P. *New Engl. J. Med.* **300,** 350–352 (1979).
2. Chance, M. L., Peters, W. & Shchory, L. *Ann. trop. Med. Parasit.* **68,** 307–316 (1974).
3. Gardener, P. J., Chance, M. L. & Peters, W. *Ann. trop. Med. Parasit.* **68,** 317–325 (1974).
4. Kilgour, V., Gardener, P. J., Godfrey, D. G. & Peters, W. *Ann. trop. Med. Parasit.* **68,** 245–246 (1974).
5. Martin, E., Simon, M. W., Schaefer, F. W. III & Muhhada, A. J. J. *Protozoology* **23,** 600–607 (1978).
6. Miles, M. A., Povoa, M. M., de Souza, A. A., Lainson, R. & Shaw, J. J. *Trans. R. Soc. trop. Med. Hyg.* **74,** 243–252 (1979).
7. Lainson, R. & Shaw, J. J. *Br. med. Bull.* **28,** 44–48 (1972).
8. Zuckerman, A. & Lainson, R. in *Parasitic Protozoa* Vol. 1 (ed. Kreier, J. P.) 57 (Academic, New York, 1977).
9. Shaw, J. J. & Lainson, R. *J. trop. Med. Hyg.* **79,** 9–13 (1976).
10. Gardener, P. J., Shchory, L., & Chance, M. L. *Ann. trop. Med. Parasit.* **71,** 147–155 (1977).
11. Camargo, M. E., Hoshino, S., Correa, N. S. & Peres, B. A. *Rev. Inst. Med. Trop. Sao Paulo* **13,** 45–50 (1971).
12. Neal, R. A. & Miles, R. A. *Rev. Inst. Med. Trop. Sao Paulo* **12,** 325–332 (1970).
13. Camargo, M. E. & Robonato, C. *Am. J. trop. Med. Hyg.* **18,** 500–505 (1969).
14. Köhler, G. & Milstein, C. *Nature* **256,** 495 (1975).
15. Kennett, R. H., Denis, K. A., Tung, A. S. & Klinman, N. R. in *Curr. Topics Microbiol. Immun.* **81,** 77 (1978).
16. Williams, A. F. *Eur. J. Immun.* **3,** 628–632 (1973).

Induction of a light requirement during seed development and its ecological consequences

E. G. Cresswell & J. P. Grime

Unit of Comparative Plant Ecology (NERC), Department of Botany, The University, Sheffield S10 2TN, UK

Germination in many arable weeds and herbaceous plants of marshland and heathland depends on exposure of the seed to unfiltered sunlight[1–4]. In contrast, seeds of many grasses, legumes and woody species exhibit no such requirement and are capable of germination in darkness[1–4]. These interspecific differences are apparent in seeds immediately after their removal from the parent plant and in theory could be due to major differences in seed physiology. Here we present evidence supporting the alternative hypothesis that the contrasting responses to light are imposed by differences in the light-filtering properties of the maternal tissues which surround the developing seeds.

Induction of a light requirement has been observed when seeds develop on plants grown under incandescent lighting[5] and when mature detached seeds receive light filtered through a leaf canopy[6–14]. These effects are the result of a decrease in the red to far-red ratio of the light incident on the seed and are mediated by the phytochrome system[15–17]. As seeds of most species are

0028-0836/81/250583—03$01.00

formed initially within green photosynthetic tissues, which only later senesce and dehisce, it is possible that, in natural conditions, a similar process of induction may take place while the developing seed is still attached to the inflorescence.

Photoconversion of phytochrome occurs only when the seed is moist[18]. A critical factor is, therefore, whether the chlorophyll content of the structures investing the seed is retained throughout maturation, or is lost before the seed is dry. In the former case the seeds will receive only filtered light with a low red/far-red ratio, whereas in the latter the tissues will lose the capacity to act as a filter and the red/far-red ratio of the light experienced by the developing seed will increase. As the seed matures its phytochrome will be arrested in a particular photostationary state, determined by the light environment present immediately before drying out. This state controls the light response of the freshly dispersed seed. Thus, seeds which have matured entirely within green tissues would have most of their phytochrome in the inactive (Pr) form, requiring a light stimulus for germination. In contrast, seeds which have been exposed to unfiltered light before drying regardless of the light quality received at earlier stages of development, would have their phytochrome in the active (Pfr) form and be able to germinate in light or in darkness.

To investigate this hypothesis, we have studied the ripening process and germination requirements of seeds removed from natural populations of species commonly occurring in Britain. In each species, seed heads providing a range of developmental stages, from newly formed to mature seed, were collected and divided into five age classes. Determination of the moisture content of the seeds in each class provided an index of maturity, and the filtering capacity of the investing structures at each stage was estimated by extraction and measurement of chlorophyll.

The investing structures were then removed in daylight from a subsample of the seeds of each maturity stage and germination tests were performed in light (16 h day, 40.0 W m^{-2} fluorescent supplemented by tungsten lighting, red/far-red 7.6) and darkness in a growth room[19] with 20 °C day and 15 °C night temperatures. The remaining intact seeds were then dried in darkness over $CaCl_2$ for ~6 weeks, after which the investing structures were removed and the germination tests repeated.

The results suggest that species conform to one of two broad types. The first, represented in Fig. 1 by *Atriplex hastata*, *Stachys sylvatica*, *Tragopogon pratensis* and *Myosotis arvensis*, includes species in which chlorophyll is retained in the investing structures throughout seed maturation. The second is composed of species such as *Arrhenatherum elatius*, *Hordeum murinum*, *Helianthemum chamaecistus* and *Vicia angustifolia* in which chlorophyll declines to a low level early in seed development.

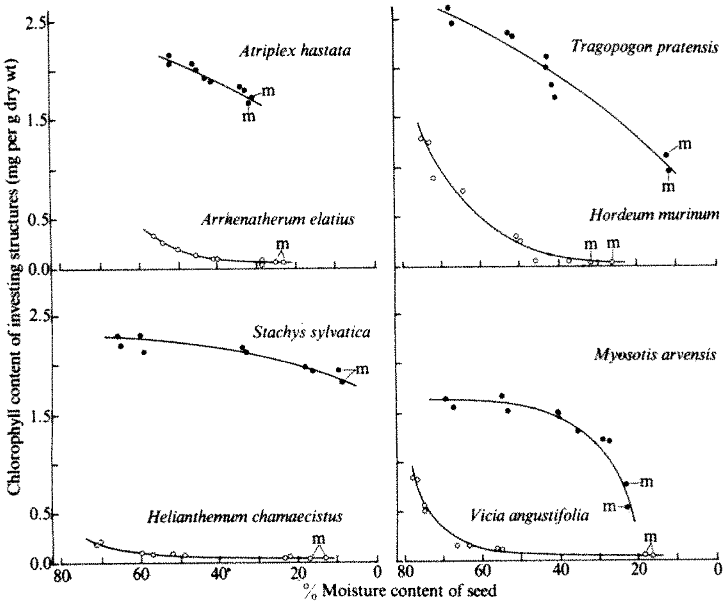


Fig. 1 Relationship between the moisture content of seeds and the chlorophyll concentration of the investing structures at various stages of inflorescence development in eight species of flowering plants. ●, Species producing mature seeds with a light requirement[4]; ○, species producing mature seeds with no light requirement[4]; m, mature seeds.

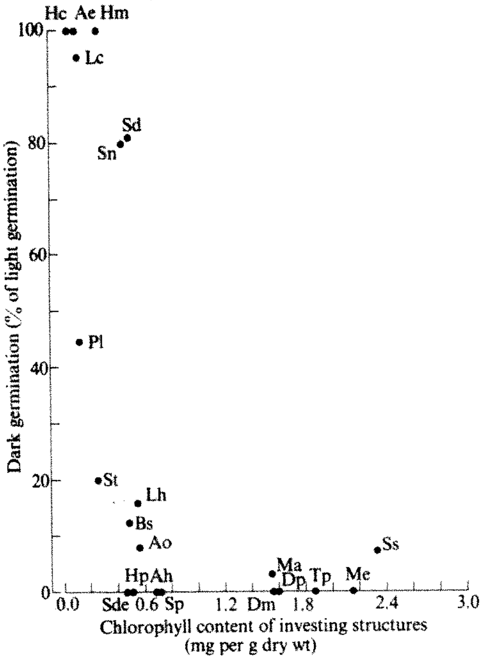


Fig. 2 Relationship between chlorophyll concentration of the investing structures and dark germination of mature seeds in various species of flowering plants. Values for chlorophyll refer to the concentration at the mid-point in the range of seed moisture contents. Ae, *Arrhenatherum elatius*; Ah, *Arabis hirsuta*; Ao, *Anthoxanthum odoratum*; Bs, *Brachypodium sylvaticum*; Dm, *Draba muralis*; Dp, *Digitalis purpurea*; Hc, *Helianthemum chamaecistus*; Hm, *Hordeum murinum*; Hp, *Hypericum perforatum*; Lc, *Lotus corniculatus*; Lh, *Leontodon hispidus*; Ma, *Myosotis arvensis*; Me, *Milium effusum*; Pl, *Plantago lanceolata*; Sd, *Silene dioica*; Sde, *Sieglingia decumbens*; Sn, *Silene nutans*; Sp, *Succisa pratensis*; Ss, *Senecio squalidus*; St, *Serratula tinctoria*; Tp, *Tragopogon pratensis*.

When the germination of mature dry-stored seeds of some of these species is examined (Fig. 2) it is evident that the retention of chlorophyll in the investing structures is strongly correlated with the presence of a light requirement.

Figure 3 shows that the light response may change during maturation and can be manipulated experimentally. *Hordeum murinum* is a species which will germinate to a comparable extent in light and darkness at any stage of maturity if the seed coats are removed in full daylight before drying the seed and conducting the germination tests. If, however, the seed is dried in darkness before removal of the investing structures, seed of the first two maturity stages will not germinate in darkness. This reveals the presence in immature seeds of *H. murinum* of an hitherto undetected phase in which a light requirement is imposed by the green parental tissues. During the normal course of maturation this requirement is lost, because the chlorophyll content declines sharply (Fig. 1) when the seed is still moist enough for its phytochrome system to respond to the resulting changes in the quality of the light transmitted to the seed. Similarly, Fig. 3 shows that the obligate light requirement of *Tragopogon pratensis* may be removed if unripe seeds are taken from their green investing bracts while still moist, and exposed to unfiltered light before undergoing the dark test. However, if these seeds are dried in darkness with the investing structures intact, the light requirement is maintained, as in the normal process of maturation within the green bracts.

These different types of response may have important ecological implications. Plants which retain chlorophyll in the investing structures throughout the maturation process tend to be the small-seeded species known to develop more or less persistent reserves of seeds buried in the soil[20-25]. The light requirement conferred on these seeds by their maturation environment may enable them to remain dormant when buried. Rather different are the larger-seeded composites, such as *Tragopogon pratensis*, *Leontodon hispidus* and *Senecio squalidus*, all of which retain chlorophyll in their bracts and exhibit strong dark inhibition of germination. Here, there is no evidence that the seeds become buried or form persistent reserves. One possible explanation, now under investigation, is that seeds released with phytochrome in the inactive form may be unusually susceptible to canopy-induced dormancy. This might allow the exploitation of vegetation gaps by delaying germination until unfiltered daylight is perceived by seeds lying on the soil surface. A similar sensitivity to canopy-induced dormancy may contribute to the delaying mechanism which allows small seeds to become buried and incorporated into persistent reserves rather than germinating soon after seed fall. An extreme example of chlorophyll retention is found in some of the ruderal species, such as *Chenopodium album* and *Atriplex hastata*, where the investing structures also maintain a high moisture content and ripe seeds occur within green fleshy tissues. In addition to its effect on the light requirement of the seed, this phenomenon may attract predation and thus facilitate seed dispersal.

Species in which chlorophyll declines at an early stage include two conspicuous groups: the large-seeded, autumn-germinating grasses and the legumes. In both, water rather than light seems to provide the environmental cue for germination. Autumn-germinating grasses, for example *Arrhenatherum elatius*, *Festuca arundinacea* and *Helictotrichon pratense*, which produce and release their seeds over the summer months, germinate more or less synchronously with the onset of wet weather in the autumn, regardless of their light environment[25-27]. The legumes, on the other hand, are held dormant by their impervious seed coats and will germinate in light or darkness once this barrier to imbibition is overcome[4,28,29].

From this study we predict that the light requirement of seeds may be strongly affected by (1) seed maturity at the time of collection, (2) removal or retention of investing structures, and (3) the light conditions during post-harvest drying. Differences in harvesting and drying procedures may account for much of the conflicting data in the literature concerning the light requirements of seeds, and differential rates of chlorophyll loss and/or seed drying within the same inflorescence may explain the germination 'polymorphism' with respect to light requirement reported for certain species[30].

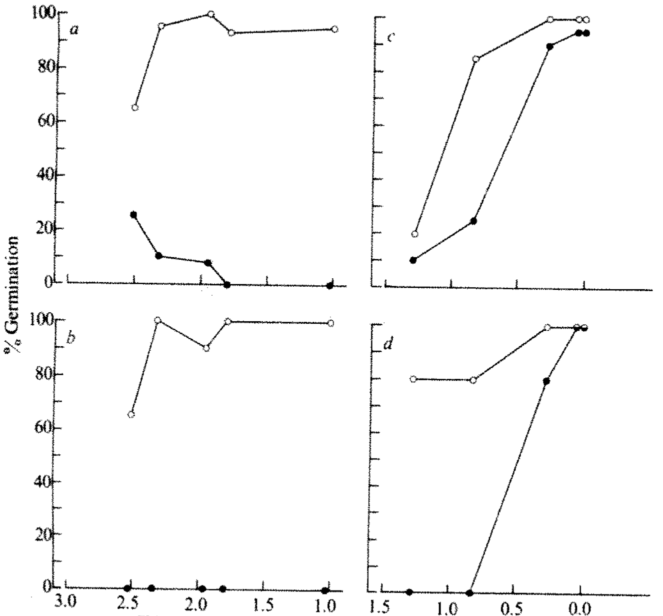


Fig. 3 Germination in light (○) and darkness (●) of seeds of *Tragopogon pratensis* (*a*, *b*) and *Hordeum murinum* (*c*, *d*). In both species, seed maturity is characterized by reference to the chlorophyll content of the investing structures. Dry storage involved placing the intact inflorescence over calcium chloride in darkness for ~6 weeks. All germination tests on fresh (upper panel) and dry-stored (lower panel) seeds were preceded by removal of the investing structures, a procedure involving ~5-min exposure to daylight.

We thank S. R. Band for his interest and assistance in the investigation. This research was supported by the Natural Environment Research Council.

Received 27 November 1980; accepted 13 March 1981.

1. Kinzel, W. *Frost und Licht als beeinflussende Krafte der Samenkeimung* (Ulmer, Stuttgart, 1920).
2. Toole, V. K. *Seed Sci. Technol.* **1**, 339–396 (1973).
3. Grime, J. P. & Jarvis, B. C. in *Light as an Ecological Factor II* (eds Bainbridge, R., Evans, G. C. & Rackham, O.) 525–532 (Blackwell, Oxford, 1975).
4. Grime, J. P. *et al. J. Ecol.* **69** (in the press).
5. McCullough, J. M. & Shropshire, W. *Pl. Cell Physiol.* **11**, 139–148 (1970).
6. Borthwick, H. A., Hendricks, S. B., Toole, M. W. & Toole, V. K. *Bot. Gaz.* **115**, 205–225 (1954).
7. Black, M. in *Dormancy and Survival* (ed. Woolhouse, H. W.) 193–217 (Cambridge University Press, 1969).
8. Pigott, C. D. *New Phytol.* **70**, 953–966 (1971).
9. Smith, H. in *Seed Ecology* (ed. Heydecker, W.) 219–231 (London, 1973).
10. Gorski, T. *Physiol. Pl.* **34**, 342–346 (1975).
11. King, T. L. *New Phytol.* **75**, 87–90 (1975).
12. Silvertown, J. W. *New Phytol.* **85**, 109–118 (1980).
13. Fenner, M. *New Phytol.* **84**, 95–102 (1980).
14. Fenner, M. *New Phytol.* **84**, 103–106 (1980).
15. Flint, L. H. & McAlister, E. D. *Smithsonian misc. Collns* **94**, 1–11 (1935).
16. Flint, L. H. & McAlister, E. D. *Smithsonian misc. Collns* **96**(2), 1–8 (1937).
17. Borthwick, H. A., Hendricks, S. B., Parker, M. & Toole, V. K. *Proc. natn. Acad. Sci. U.S.A.* **38**, 662–666 (1952).
18. Evenari, M. *Handb. PflPhysiol.* **15**, 804–847 (1965).
19. Rorison, I. H. *New Phytol.* **63**, 358–362 (1964).
20. Brenchley, W. E. *J. agric. Sci.* **9**, 1–31 (1918).
21. Chippindale, H. G. & Milton, W. E. J. *J. Ecol.* **22**, 508–531 (1934).
22. Milton, W. E. J. *J. Ecol.* **27**, 149–159 (1939).
23. Champness, S. S. & Morris, K. *J. Ecol.* **36**, 149–173 (1948).
24. Wesson, G. & Wareing, P. F. *J. exp. Bot.* **20**, 401–413 (1969).
25. Thompson, K. & Grime, J. P. *J. Ecol.* **67**, 893–921 (1979).
26. Chancellor, R. J. *Proc. 9th Br. Weed Control Conf.*, 1129–1135 (1968).
27. Davison, A. W. *J. Ecol.* **59**, 493–506 (1971).
28. Quinliven, B. J. *Aust. J. agric. Res.* **12**, 1009–1022 (1961).
29. Ballard, L. A. T. *Seed Sci. Technol.* **1**, 285–303 (1973).
30. Cavers, P. B. & Harper, J. L. *J. Ecol.* **54**, 307–382 (1966).

Depression of net photosynthesis in *Vicia faba* L. exposed to sulphur dioxide and ozone

D. P. Ormrod, V. J. Black & M. H. Unsworth

Department of Physiology and Environmental Studies, University of Nottingham, Sutton Bonington, Loughborough LE12 5RD, UK

In most laboratory investigations of plant responses to air pollution, plants are exposed to a single pollutant gas. However, in rural atmospheres, pollutants seldom occur singly; for example, in central England, where summertime mean concentrations of sulphur dioxide are typically 10–30 p.p.b. (parts in 10^9 by volume) depending on the site, there are 10–30 days each summer when concentrations of ozone generated photochemically exceed 80 p.p.b. for at least 1 h (ref. 1). Maximum hourly ozone concentrations of 130–260 p.p.b. have been reported[1-3]. Similarly, maximum hourly SO_2 concentrations in rural central England exceed 50 p.p.b. for a few hours each month. We have previously shown[4] that net photosynthetic rates P_N of field beans, *Vicia faba* L. c.v. Dylan, were depressed, in comparison with rates in clean air, on exposure to SO_2 concentrations as low as 40 p.p.b. Responses of photosynthesis to mixtures of SO_2 and O_3 seem to have been studied only at higher concentrations atypical of the natural environment[5,6]. We report here that the addition of ozone in the concentration range 60–150 p.p.b. to atmospheres containing 40 p.p.b. SO_2 led to decreases in P_N of *V. faba* L. larger than those observed in either gas alone.

0028–0836/81/250585—02$01.00

Three-week old plants of *V. faba* L. c.v. Dylan, grown in a controlled environment were exposed for 4 h in chambers described fully elsewhere[7] at 22 ± 2 °C, 280 W m^{-2} photosynthetically active radiation, and 0.9 ± 0.2 kPa vapour pressure deficit. Air flowed through the chambers at approximately 20 l min^{-1} and internal fans minimized boundary layer resistance.

Photosynthetic rates of plants exposed to pollutants were compared with those of 'control' plants in an identical chamber containing clean air. Before exposure, plants were acclimatized for 24 h with clean air in both chambers. During the period immediately before exposure to pollutants, rates of photosynthesis were compared between chambers; typical rates were 2 g CO_2 m^{-2} h^{-1}. Then air containing either 40 p.p.b. SO_2 or O_3 or $O_3 + 40$ p.p.b. SO_2 was supplied to one chamber and P_N was monitored in both chambers. There were 18 experiments with O_3 and 21 with $O_3 + SO_2$; in each experiment three plants were exposed in one pot in each chamber.

When either O_3 or $O_3 + SO_2$ was added, P_N decreased. The maximum depression of P_N during exposure increased with increasing O_3 concentration. The change δP_{NT} between P_N of treated plants on the day before exposure and the minimum P_N during exposure was assumed to be related linearly to the change δP_{NC} between photosynthetic rates of control plants on the two days, and to O_3 concentration. Least-squares regression analysis was used to fit the expression

$$\delta P_{NT} = a\delta P_{NC} + b[O_3] + c \qquad (1)$$

where a, b and c are constants, to the measurements with O_3 alone and with $O_3 + SO_2$. The value of a, which is a measure of the influence of environmentally and genetically determined responses of P_N, was held constant between the two sets of observations. The value of b is the rate of decrease of P_N with O_3 concentration [units: g CO_2 m^{-2} h^{-1} (p.p.b. $O_3)^{-1}$]. The value of c [units: g CO_2 m^{-2} h^{-1}] gives the depression of P_N at zero O_3 concentration. For the treatment with O_3 alone, c therefore indicates the magnitude of residual variation in P_N not accounted for by variation in the control. For the treatment with $O_3 + SO_2$, c also gives the depression in P_N caused by SO_2 alone.

Values of b and c differed between the two treatments and an analysis of variance was used to establish the statistical significance of the differences. Table 1 shows the values of a, b and c for the $O_3 + SO_2$ treatment, and the differences from values derived for O_3 alone. The differences are highly significant ($P < 0.01$).

The regression coefficients indicate that photosynthesis was significantly depressed by 40 p.p.b. SO_2 alone (a point reported previously[4]), but the addition of increasing concentrations of O_3 to a background of 40 p.p.b. SO_2 depressed photosynthesis further. The rate of decrease in P_N with increasing O_3 concentration was larger in O_3 alone than in $O_3 + 40$ p.p.b. SO_2.

Figure 1 illustrates the regression equations for the case $\delta P_{NC} = 0$. At low O_3 concentrations, the depression of P_N was greater (δP_{NT} more negative) in $O_3 + SO_2$ than in O_3 alone. At ozone concentrations above about 150 p.p.b., the two treatments produced similar depressions in P_N. It may be significant that the threshold for visible injury by O_3 in these experiments was at about this concentration.

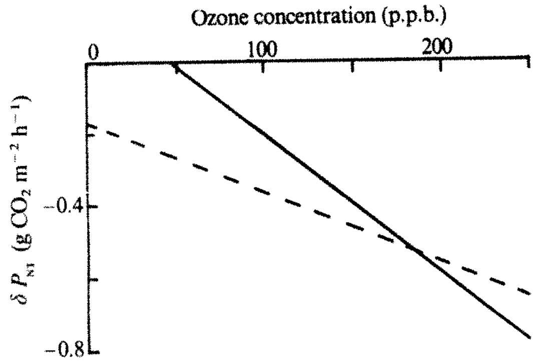


Fig. 1 Regressions of the depression of net photosynthesis δP_{NT} against ozone concentration for *V. faba* plants exposed to O_3 alone (———) or to $O_3 + 40$ p.p.b. SO_2 (– – –). δP_{NT} = (P_N during exposure) − (P_N before exposure).

Table 1 Coefficients a, b and c with standard errors in equation (1)

$$\delta P_{NT} = a\delta P_{NC} + b[O_3] + c$$

	$O_3 + 40$ p.p.b. SO_2	*Difference Δ*
a	0.393 ± 0.224	—
b	−0.0019 ± 0.0004	−0.0020 ± 0.0007
c	−0.170 ± 0.068	0.345 ± 0.110

δP_N is the difference in net photosynthetic rate during exposure to pollutants and before exposure. Suffices T and C refer to plants exposed to pollutants and control plants respectively. Values of a, b and c for $O_3 + SO_2$ exposures are shown; differences Δ between b and c for the O_3 and $O_3 + SO_2$ exposures are also shown, such that $b(O_3) = b(O_3 + SO_2) + \Delta$ and so on. Units: a is non-dimensional; b, g CO_2 m^{-2} h^{-1} (p.p.b. $O_3)^{-1}$; c, g CO_2 m^{-2} h^{-1}.

These results suggest that photosynthesis which is depressed by SO_2 at concentrations commonly occurring in Britain would be depressed further by the superposition of ozone generated photochemically. This effect would be most important near urban areas with sources of appropriate pollutants and precursors.

On the basis of these results we propose the following general mechanism for responses to mixtures of pollutant gases. At low concentrations of pollutants the injury done by one pollutant may be insufficient to influence an independent response to a second pollutant. High concentrations of one pollutant may lead to substantial injury which masks or over-rides any response to a second pollutant. Interactions between pollutants and subsequent synergistic responses seem most likely to be observed at concentrations between these extremes.

This work was supported by NERC and SRC. We thank H. L. Back for advice on statistical analysis.

Received 30 September 1980; accepted 10 April 1981.

1. Martin, A. & Barber, F. R. *Atmos. Envir.* (in the press).
2. Bell, J. N. B. & Cox, R. A. *Envir. Pollut.* **8,** 163–170 (1975).
3. Apling, A. F. *et al. Nature* **269,** 569–573 (1977).
4. Black, V. J. & Unsworth, M. H. *J. exp. Bot.* **30,** 473–480 (1979).
5. Bennett, J. H. & Hill, A. C. in *Air Pollution Effects on Plant Growth* (ed. Dugger, M.) *ACS Symp. Ser.* **3,** 115–127 (American Chemical Society Washington, 1974).
6. Carlson, R. W. *Envir. Pollut.* **18,** 159–170 (1979).
7. Black, V. J. & Unsworth, M. H. *J. exp. Bot.* **30,** 81–88 (1979).

Gene transfer in *Nicotiana rustica* using irradiated pollen

J. L. Jinks, P. D. S. Caligari* & N. R. Ingram

Department of Genetics, University of Birmingham,
Birmingham B15 2TT, UK

In spite of reports of widespread matromorphy—the occurrence of diploid plants whose genomes are exclusively maternal in origin—in *Nicotiana*[1] and in other genera[2,3], attempts to induce it in pure-breeding material of *Nicotiana rustica* have met with little success[4,5]. When heavily irradiated pollen from one variety is used as the inducer on a contrasting variety acting as maternal parent, the progeny, while showing a greater resemblance to the maternal parent, are not true matromorphs because they usually have at least one characteristic of the paternal parent. At the highest dose of γ-irradiation, 20 krad, the progeny usually

---

* Present address: Scottish Crop Research Institute, Pentlandfield, Roslin, Midlothian EH25 9RF, UK.

0028-0836/81/250586—03$01.00

Table 1 Five characters recorded for the maternal (V27) and paternal (V12) parents, the F_2 of the cross between these, and the corresponding M_2 obtained by self-pollination of M_1 plants

	Plant colour Yellow/green	Inflorescence Mop/non-mop	Height at flowering time (cm)	Height 10 weeks after planting (cm)	Final height (cm)
V27	25:0	25:0	77.02±1.18	110.96±1.02	120.60±1.10
V12	0:25	0:25	126.46±2.64	163.84±1.60	170.60±2.30
F_2	2:98	3:97	121.39±2.08	173.68±1.94	180.97±2.30
M_2	16:81	23:74	93.77±2.85	135.60±3.27	148.27±3.43
Selection of M_2 plants					
Yellow plants (16/97)	–	10:6	85.06±6.07	115.25±8.23	128.63±8.31
Yellow plants + mop head (10/16)	–	–	79.00±5.65	103.60±9.58	121.70±11.89
Yellow plants + mop head with black pigment (6/10)	–	–	81.17±6.78	114.17±12.65	131.17±12.91

The M_1 plants resulted from pollination of V27 with pollen of V12 that had been γ-irradiated (20 krad). Selections were made from the M_2 population on the basis of recessive yellow plant colour and mop-head inflorescence (typical of V27), and the dominant black-pigmented ovary (typical of V12). Height values are mean ± s.e.m.

acquire from the paternal parent only one or a small sample of the many heritable characteristics by which the two varieties differ. By selecting within these progenies we have shown here that lines can be isolated with the characteristics of the pure-breeding maternal variety but with the exception of a specific characteristic transferred from the paternal variety. To achieve the same end by conventional means would require a combination of many generations of recurrent backcrossing and selection.

We have worked with the varieties V12 and V27 of *N. rustica*. For several genetic characteristics (for example, colour) V12 is dominant over V27 and was therefore chosen as the paternal parent. Irradiated pollen (10–20 krad) of V12 was used to pollinate the pure-breeding V27 to yield progeny M_1, from which progeny M_2 were obtained by random self-pollination. We have compared the characteristics of M_2 with those of the progeny F_2 obtained by crossing the two varieties using unirradiated pollen. The M_2 progenies resembled V27 more than did F_2 progenies from the same cross made with unirradiated pollen[5], the similarity to V27 increasing with the three doses used. However, although all the M_2 families produced from the highest dose closely resembled the initial maternal parent, V27, there was significant phenotypic heterogeneity among them. We concluded that the most plausible explanation was the effective transfer of incomplete, random parts of the genome of the irradiated parent into the M_1 and M_2 progenies whose genomes were, therefore, largely derived from V27, the maternal parent of the M_1 crosses. Further segregational data and cytological examinations are required to test this explanation but none of the usual signs of gross chromosomal abnormalities are present in the M_2.

Because the M_2 families resulting from the highest dose of radiation showed not only the greatest resemblance to V27 but also significant heterogeneity in their phenotypes, there are obviously opportunities for selection among them. It would, for example, be of particular practical significance if we could demonstrate the presence of individuals identical to V27 except for a single characteristic which they have acquired from the pollen parent, V12: that is, if we could demonstrate the practicability of transferring single desirable characters from one genotype into another without the simultaneous acquisition of other less desirable characters which would then have to be bred out by a combination of recurrent backcrossing and selection.

Table 1 gives the results of five characters recorded for the M_2 progenies derived from a cross involving V27 (maternal parent) and pollen of V12 which had received 20 krad of γ rays. Two of these characters, plant colour and inflorescence morphology, display the effects of major gene differences, thus the results are given as segregation ratios. The other three characters, height at flowering time, height 10 weeks after planting and final height, are continuously variable and are therefore expressed as means (cm) ± s.e.m. The corresponding results for V27, V12 and their F_2 are also given. For every character the F_2 is more like V12 than V27, that is, V12 has mainly the dominant alleles for these characters.

If irradiation of the V12 pollen had no effect, the results for the M_2 and F_2 generations would differ only by sampling variation, but this was not the case. For each character, the M_2 shows a greater resemblance to the recessive maternal parent, V27. However, the continuously varying characters are correlated, thus the information they provide is not completely independent[5].

Plant colour is the first character that can be scored during development thus it provides the first criterion for selecting maternal phenotypes. Of the 97 M_2 plants, 16 showed the recessive yellow colour of V27. During the development of these 16 plants, the other characters were scored (see Table 1). Among these plants a high proportion (10 out of 16) also expressed the other recessive major gene-controlled character of V27, mop-head inflorescence. This proportion is surprisingly high as there were no plants with the combination of yellow colour and mop-head inflorescence in the equivalent F_2.

M_2 plants selected for yellow plant colour also had mean heights that were closer to those of the maternal parent, V27, than those of the unselected M_2. If from among the yellow plants of the M_2, those which are also mop heads are selected, they have means which are even closer to those of V27 for height at flowering time and final height. However, the mean height 10 weeks after planting is below even the low mean of V27 because of two particularly low scoring plants in the small sample of 10.

As our intention was to examine the possibility of transferring a paternal character from V12 into the maternal genotype of V27, we next selected individuals from among those 10 plants which expressed the appropriate paternal characteristic: the remaining major gene-controlled difference, namely, the black anthocyanin pigmentation present in the ovary of V12 and absent in V27. Of the 10 plants, 6 had this pigmentation and were therefore of the desired type. Thus from the original 97 M_2 plants it was possible to select six with the maternal phenotype for two major gene controlled characters and paternal phenotype for the third. No plant with this phenotype was present in a sample of 100 F_2 of the same cross.

The means of the quantitative characters (see Table 1) for these six selected plants, although very close to those of the maternal parent, V27, do show a slight deviation when compared with the 10 plants from which they were chosen. This suggests that some factor, although very minor, has been transferred with the pigment gene or alternatively that the latter has a pleiotropic effect on the quantitative characters.

In an earlier experiment (ref. 4 and D. S. Virk and G. S. Pooni, unpublished data) in which only plants selected for yellow colour, mop-head inflorescence and black anthocyanin pigmentation of the ovary were taken beyond the M_1 generation, the presence versus the absence of the pigment segregated in over half of the progenies produced by self-pollinating pigmented plants in the M_2, M_3 and M_4 generations. The progenies produced by self-pollinating the remaining pigmented plants and all the recessive non-black segregants in the M_2 bred true in all subsequent generations. Therefore, we can readily extract families from the M_2 which are pure-breeding for the recessive maternal traits, yellow plant colour and mop-head inflorescence and the dominant paternal black ovary, although M_3 progenies have to be raised to distinguish between plants which are homozygous, and therefore pure-breeding for black ovary, from those which are heterozygous and therefore still segregating for the recessive non-black ovary.

Because of the greater clarity of the results, we have concentrated in these selection studies on the major gene controlled characters. The same questions, however, can be posed for the quantitative differences between V27 and V12. For example, can plants be selected which are identical to V27 except in respect of one quantitative character for which they are more like V12? Of the 97 M_2 plants, 4 were identical to V27 for all the major gene controlled characters. One of these had a final height of 158 cm, which is more like that of V12, while the other quantitative characters were more like V27. Similarly in the earlier experiment (ref. 4 and D. S. Virk and G. S. Pooni, unpublished data) some of the 120 families which had all the major gene-controlled characteristics of V27 were more like V12 for one of the seven quantitative characters recorded.

It is premature to offer explanations for the similarities and differences between our results and those of Pandey[6,7]. At this early stage it is more important to establish a working model that allows the routine transfer of one or more characteristics of a paternal parent to progeny that are overwhelmingly like the maternal parent and at the same time will allow the more fundamental questions about mechanisms to be investigated. By selecting in the M_2 in our model system we can transfer specific paternal characteristics into a maternal background with a speed and reliability that would be difficult to match by conventional breeding. This work has already prepared the way for similar transfers between more distantly related paternal and maternal parents which do not produce successful sexual hybrids.

We acknowledge the financial support of the Agricultural Research Council and a postgraduate research studentship from the SRC to N.R.I. We thank Professor Sir Kenneth Mather for his advice and encouragement and Dr D. S. Virk and G. S. Pooni for allowing us to refer to their unpublished data.

Received 10 January; accepted 26 March 1981.

1. Goodspeed, T. H. *The Genus* Nicotiana (Chronica Botanica, Massachusetts, 1954).
2. Eenink, A. H. *Euphytica* **23**, 429–433 (1974).
3. Eenink, A. H. *Euphytica* **23**, 435–445 (1974).
4. Virk, D. S., Dhahi, S. J. & Brumpton, R. J. *Heredity* **39**, 287–295 (1977).
5. Caligari, P. D. S., Ingram, N. R. & Jinks, J. L. *Heredity* (in the press).
6. Pandey, K. K. *Nature* **256**, 310–313 (1975).
7. Pandey, K. K. *Genetics* **49**, 56–69 (1978).

Visual cortex of an anthropoid ape

François Vital-Durand* & Colin Blakemore†

*Laboratoire de Neuropsychologie Expérimentale, INSERM,
Unité 94, 16 Avenue du Doyen Lépine, 69500 Bron, France
†University Laboratory of Physiology, Parks Road,
Oxford OX1 3PT, UK

Marg *et al.*[1] recorded activity from an unspecified region of the human occipital cortex but the properties of neurones there were very different from those in cat or monkey striate cortex. Work on man's closest relatives, the apes[2], might give a better indication of the way in which visual information is analysed in the human cortex. We had an unusual opportunity to study an old, non-reproductive, female black gibbon (*Hylobates concolor*, a lesser Asian ape from the forests of south-east Asia). The properties of neurones in striate and pre-striate cortex in this animal suggest that the visual areas of anthropoid apes are similar in their organization to those of monkeys[3–5].

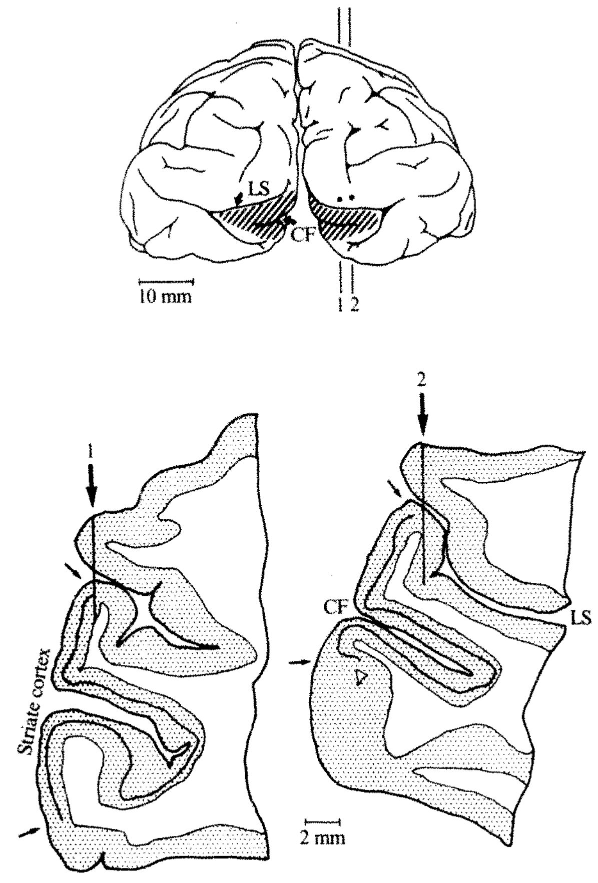


Fig. 1 The top of the figure is a view from behind of the cerebral hemispheres of the gibbon. The exposed part of the striate cortex (area 17) around the calcarine fissure (CF) is shown striped. Filled circles represent the points of entry of the two vertical penetrations, just anterior to the lunate sulcus (LS) in the right hemisphere. The diagrams at the bottom are of sections in the parasagittal planes marked 1 and 2 on the drawing of the brain. The grey matter is stippled and area 17, whose anterior and posterior boundaries are marked with small arrows, is distinguished by the stria of Gennari, shown as a line in the grey matter. Both electrode tracks passed through pre-lunate cortex and across the lunate sulcus: track 1 then re-entered grey matter just anterior to the 17/18 border and ran into the deeper layers of striate cortex; track 2 went down the posterior bank of the lunate sulcus.

The gibbon was premedicated with corticosteroid, tranquillized with ketamine and prepared for an electrophysiological experiment under intravenous (i.v.) steroid anaesthetic (Althesin, Glaxo). During recording, anaesthesia and paralysis were maintained with a continuous i.v. infusion of Nembutal and Flaxedil in glucose–saline solution. The animal was artificially ventilated with air, maintaining end-expiratory CO_2 at ~4.5%. The general experimental techniques were similar to those used previously in monkeys[6]. The eyes, protected with contact lenses, were brought to focus, by means of additional spectacle lenses, on a tangent screen at a distance of 114 cm, where we back-projected patterns to investigate the receptive fields of cortical neurones.

The gross disposition of the striate area in the occipital cortex of the gibbon is similar to that of the chimpanzee[7] where visually evoked potentials have been recorded (C. N. Woolsey and C. Welt, personal communication). Much of the primary visual cortex is buried in the calcarine fissure and only a small strip of striate cortex is exposed on the posterior pole of the hemispheres, around the lips of the fissure (Fig. 1).

We made two, long stereotaxically-vertical penetrations with a glass-insulated tungsten microelectrode[8] through the occipital cortex of the right hemisphere, starting in pre-striate cortex,

0028-0836/81/250588—03$01.00

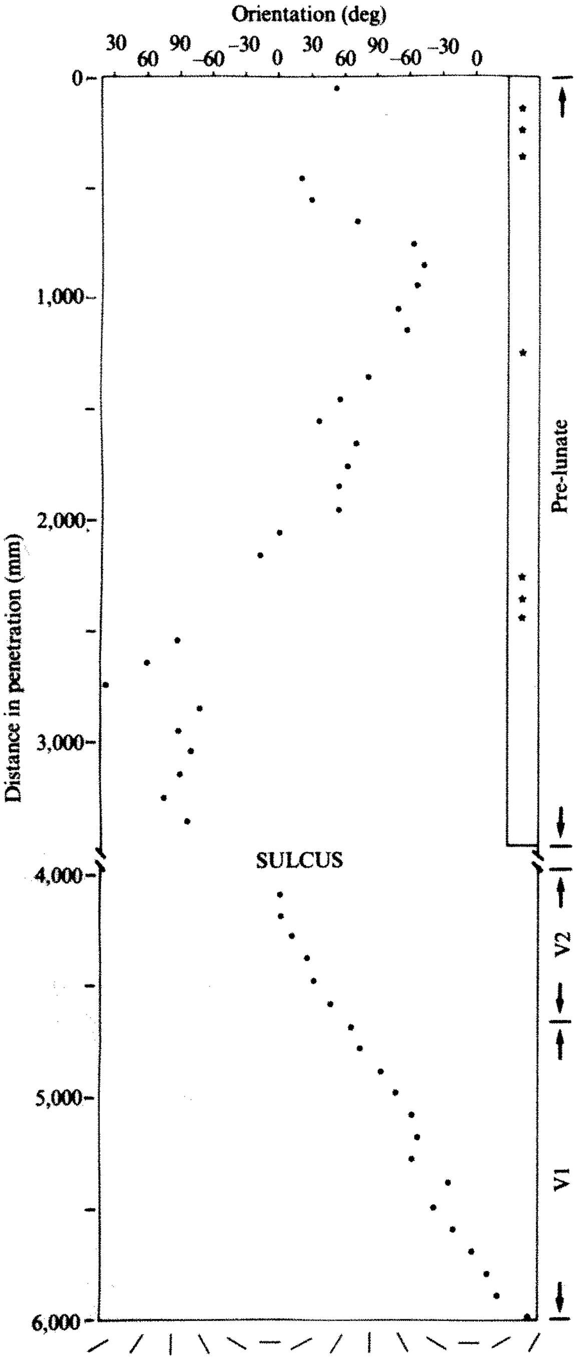


Fig. 2 Properties of neurones recorded in track 1 (see Fig. 1) are reconstructed as a function of distance along the penetration. Orientation-selective cells are plotted as filled circles, the abscissa being the angle of a bar or edge that produced the best response. Non-oriented cells are plotted as stars in the column at the right. The upper part of the track was in pre-lunate cortex, it then crossed the sulcus, entered V2 and passed into V1 (striate cortex). Note the regular sequential shifts in preferred orientation from cell to cell in V1 and V2.

anterior to the lunate sulcus. In both penetrations (see Fig. 1) the electrode passed through this pre-lunate area, across the lunate sulcus and into the anterior lip of the pre-calcarine operculum, just ahead of the cytoarchitectonic border between areas 17 and 18. Track 1 then ran into area 17, while track 2 passed into the posterior bank of the lunate sulcus, which would be occupied by the second visual area in monkeys. All 79 neurones isolated during these penetrations were visually responsive and their receptive fields lay within 4° of the horizontal meridian.

Primary visual cortex (V1): we studied only 14 cells in the striate cortex but nevertheless we were struck by their similarity to neurones in area 17 of the monkey. All were selective for the orientation of a bar or edge; they could be classified as simple or complex[9] and some had a preference for short lines. Preferred orientation shifted progressively as the electrode passed obliquely through the cortex (Fig. 2) in the 'columnar' pattern typical of monkeys[10]. Receptive fields were small (average size 0.79 deg^2) and they were all close to the vertical meridian, just above the foveal projection. All but one of the neurones were binocularly driven (Fig. 3). Unfortunately the electrode entered the striate cortex just below layer IVc, thus we had no opportunity to determine whether this layer contained non-oriented, monocularly driven cells arranged in ocular dominance clusters: certainly in the chimpanzee[11], as in monkeys[12,13], the left- and right-eye afferent axons terminate in layer IVc in alternating stripes.

Second visual area (V2): the vertical meridian is represented at the border between area 17 behind, and 18 in front. The retinotopic map is mirror-reversed in the portion of area 18 that runs forward from the border into the posterior bank of the lunate sulcus: the receptive fields of cells encountered in this region, presumably homologous to V2 of the monkey[14], moved from the vertical meridian into the contralateral field as the electrode moved down the posterior bank away from the 17/18 border in track R2. The receptive fields were a little larger than in V1 (mean size 1.41 deg^2) but all but one of the 21 cells isolated in V2 were orientation-selective, there was again evidence of a 'columnar' pattern (Fig. 2) and only one cell was not binocularly driven (Fig. 3).

Pre-lunate visual area: 44 cells were recorded in the pre-lunate area. In both tracks, as the electrode moved down, the receptive fields shifted in towards the vertical meridian from ~11 deg left to ~6 deg left. The receptive fields of cells found in the first penetration were very close to the horizontal meridian whereas those for the more lateral track were slightly higher in the visual field. This suggests that there is a topographically organized pre-lunate visual area, for which the representation of the vertical meridian lies somewhere in the anterior bank of the lunate sulcus. This presumably means that this area is not homologous to V3 in the rhesus monkey, as V3 joins V2 at the

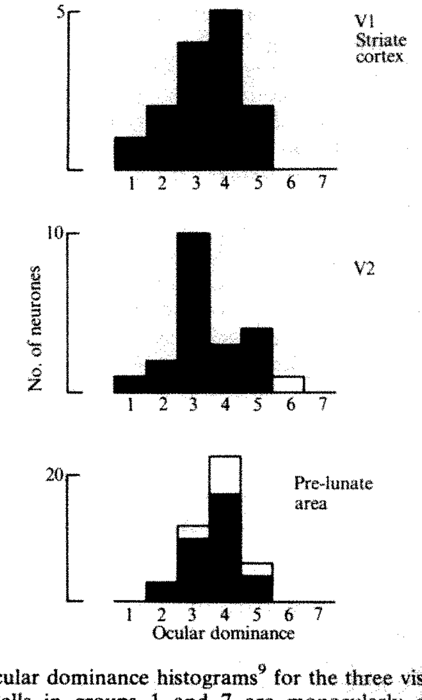


Fig. 3 Ocular dominance histograms[9] for the three visual areas studied. Cells in groups 1 and 7 are monocularly driven by contralateral and ipsilateral eyes respectively. All other groups represent binocular neurones, with a spectrum of dominance from strong dominance by the contralateral eye (group 2), through equal excitability (group 4), to strong dominance by the ipsilateral eye (group 6). Orientation-selective cells are represented by solid blocks, non-oriented cells by open blocks.

representation of the horizontal meridian[14], but that it may be equivalent to area V3A (ref. 14). This, in turn, implies that there is at least one other visual area between V2 and this pre-lunate area.

The receptive fields of pre-lunate cells were quite large (average size 8.43 deg^2), 34 of these (77%) were selective for orientation but the remainder were neither orientation- nor direction-selective. Of the 34 cells with an orientation preference, 18 had some degree of direction selectivity (a slightly higher proportion than in V1 or V2) and each cell encountered was binocularly driven (Fig. 3). Only one cell had a clear colour preference when tested with broad-band coloured stimuli; thus it is unlikely that this area is homologous with the pre-striate V4 complex to which Zeki[15] ascribes the analysis of chromatic information. Indeed we were struck by how few cells in any area had an obvious colour preference.

Our results suggest that areas V1 and V2 in anthropoid apes are organized much as in monkeys and that there is probably a patchwork of further pre-striate areas, stretching forwards towards the parietal and temporal lobes. The close genetic similarity of gibbon to man makes it likely that the human visual areas are similarly arranged.

We thank Dr Vre J. Rinjard for supplying the gibbon, which was being disposed of by the Museum d'Histoire Naturelle de Paris. The work was supported by grants from INSERM and MRC.

Received 28 January; accepted 9 April 1981.

1. Marg, E., Adams, J. E. & Rutkin, B. *Experientia* **24**, 348–350 (1968).
2. Bruce, E. J. & Ayala, F. J. *Nature* **276**, 264–265 (1978).
3. Zeki, S. M. in *Essays on the Nervous System* (eds Bellairs, R. & Gray, E. G.) 327–343 (Clarendon, Oxford, 1974).
4. Allman, J. *Prog. Psychobiol. physiol. Psychol.* **7**, 1–53 (1977).
5. Van Essen, D. C. *A. Rev. Neurosci.* **2**, 227–263 (1979).
6. Blakemore, C., Garey, L. J. & Vital-Durand, F. *J. Physiol., Lond.* **283**, 223–262 (1978).
7. Polyak, S. *The Vertebrate Visual System* (University of Chicago Press, 1957).
8. Merrill, E. G. & Ainsworth, A. *Med. biol. Engng* **10**, 662–672 (1972).
9. Hubel, D. H. & Wiesel, T. N. *J. Physiol., Lond.* **195**, 215–243 (1968).
10. Hubel, D. H. & Wiesel, T. N. *J. comp. Neurol.* **158**, 267–293 (1974).
11. Tigges, J. & Tigges, M. *Brain Res.* **166**, 386–390 (1979).
12. Wiesel, T. N., Hubel, D. H. & Lam, D. M. K. *Brain Res.* **79**, 273–279 (1974).
13. Hendrickson, A. E., Wilson, J. R. & Ogren, M. P. *J. comp. Neurol.* **182**, 123–136 (1978).
14. Van Essen, D. C. & Zeki, S. M. *J. Physiol., Lond.* **277**, 193–226 (1978).
15. Zeki, S. *Nature* **284**, 412–418 (1980).

Thrombolysis by human tissue plasminogen activator and urokinase in rabbits with experimental pulmonary embolus

O. Matsuo, D. C. Rijken & D. Collen

Center for Thrombosis and Vascular Research, Department of Medical Research, University of Leuven, Belgium

Thromboembolism is a frequent and often lethal complication of many medical diseases and surgical procedures. Streptokinase, a bacterial non-enzymatic protein, and urokinase, an enzyme obtained from human urine, are both potent activators of the human fibrinolytic system, and have been extensively investigated as a means of medical treatment of this condition. Although it is well established that these agents can clear obstructed vessels in roughly 50% of cases, the fact that they induce relatively extensive systemic fibrinogen breakdown and a serious haemorrhagic diathesis has prevented their general application for thrombolysis. The physiological plasminogen activator in blood (blood activator), which is probably released from the vascular wall (vascular activator) and is identical or at least very similar to the activator extracted from human organs (tissue activator), differs from urokinase (for references see ref. 1). This activator has a markedly higher fibrinolytic to fibrinogenolytic ratio than urokinase *in vitro*[1] and might therefore constitute a better thrombolytic agent. We have developed a purification method for tissue plasminogen activator from a human melanoma cell line which enables us to produce milligramme quantities on a laboratory bench scale[2] and we have now compared its thrombolytic effect with that of urokinase in an experimental animal model. We have found that tissue plasminogen activator causes thrombolysis at lower doses than urokinase, without extensive plasminogen activation in the circulating blood and without haemostatic breakdown.

The fibrinolytic activities of tissue plasminogen activator and urokinase were determined on plasminogen-enriched bovine fibrin plates[3] and expressed in International Units (IU) using the WHO 1st International Reference Preparation of urokinase as described previously[4]. The specific activities of both enzymes were very similar to those described elsewhere[2]. Human fibrinogen was prepared as described by Blombäck and Blombäck[5] and labelled with ^{125}I according to McFarlane[6].

An experimental thrombus was produced in a polyethylene tube of 3 mm internal diameter by mixing citrated whole human blood (0.5 ml) with 8 µl of ^{125}I-labelled human fibrinogen (~10^6 c.p.m.), 50 µl of 25 mM $CaCl_2$ and 2 µl of human thrombin (final concentration 1 NIH unit per ml). After incubation at room temperature for 15 min, 1.5 cm length of the tube (~0.10 ml thrombus) was cut out of the middle position of the tube. The thrombus was poured into a Petri dish, washed repeatedly with 0.15 M NaCl and its ^{125}I content measured. The radioactive thrombus was then aspirated in a silicon catheter for injection.

New Zealand White rabbits were anaesthesized with 0.5 ml Hypnorm and 0.3 ml pentobarbital. A jugular vein was dissected free, cut over a distance of 3 mm and, after injection of the thrombus, carefully sutured to restore the blood flow. The activator infusions (tissue plasminogen activator or urokinase) were given intravenously into a contralateral marginal ear vein as a bolus injection of 10% of the total dose followed by a continuous infusion over 6 or 12 h as indicated; blood samples were collected from a femoral vein catheter on citrate at regular

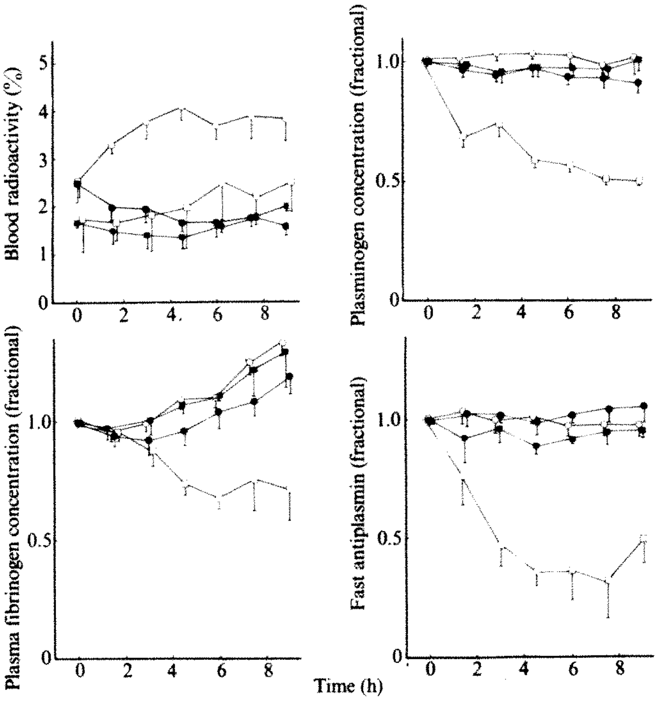


Fig. 1 Changes in the main blood parameters during infusion of tissue plasminogen activator or urokinase in rabbits. ■, Control animals infused with saline; ●, 100,000 units of urokinase in five rabbits; □, 1,000,000 units of urokinase in three rabbits; ○, 35,000 units of tissue plasminogen activator over 6 h in six rabbits and 70,000 units over 12 h in three rabbits. The data represent mean values ±s.d. Blood radioactivity was expressed as per cent of the injected radioactivity present in the total blood volume.

0028–0836/81/250590—02$01.00

Table 1 Extent of thrombolysis and isotope recovery

Group	n	% Thrombolysis	% Recovery
Saline 9 h	4	3.2 ± 1.5	102.0 ± 2.4
Saline 24 h	6	5.8 ± 3.2	100.3 ± 4.4
UK 100,000	5	4.0 ± 1.2	103.3 ± 1.1
UK 1,000,000	3	11.5 ± 5.4	97.7 ± 2.9
TA 35,000	6	16.5 ± 2.7	100.4 ± 3.3
TA 70,000	3	23.1 ± 7.4	97.0 ± 0.8

The extent of thrombolysis was determined as the difference in radioactivity between the injected and recovered clots. The isotope recovery was measured as the sum of the radioactivity present in the recovered thrombus, the blood, the urine and the lungs and expressed as per cent of the injected tracer. UK, urokinase; TA, tissue plasminogen activator; n, number of experimental animals. The data represent means ± s.d.

time intervals for up to 24 h. The animals were killed 9 or 24 h after the start of the infusion, the thorax was opened, the lungs and heart removed together and the thrombus localized with the use of an isotope monitor. The lungs were then carefully dissected, the pulmonary artery incised and the remaining radioactive thrombus recovered. The degree of thrombolysis was calculated as the difference in ^{125}I content of the injected and recovered thrombus. In all but three instances the remaining thrombus was recovered as a single fragment, tightly lodged in a major branch of the pulmonary artery.

Figure 1 represents changes in the main blood parameters during the first 9 h of the infusion. The radioactivity in the blood rose to ~2% of the amount injected in samples taken after 15 min. In the control groups without activator infusion and in the urokinase groups the blood radioactivity remained at this level throughout the experiment, whereas infusion of tissue plasminogen activator (35,000 IU over 6 h or 70,000 IU over 12 h) resulted in a significant rise in blood radioactivity (from 2.2% to a maximum of 4.1%). The fibrinogen level[5] rose progressively in controls and in those given 10,000 IU urokinase tissue plasminogen activator (infused over 6 h), probably as a result of an acute phase reaction. On infusion of 1,000,000 IU urokinase over 6 h the fibrinogen level fell to 70%. The plasminogen[7] and fast antiplasmin[8] levels did not change significantly in the control, 100,000 IU urokinase and tissue plasminogen activator groups, but fell to 50 and 30%, respectively, in the group infused with 1,000,000 IU urokinase.

The extent of thrombolysis and the isotope balance (sum of radioactivity in thrombus, blood, urine and lungs) are represented in Table 1. Significant thrombolysis occurred within 9 h when 1,000,000 IU of urokinase ($P < 0.05$) were infused over 6 h but not with 100,000 IU of urokinase. Infusion of 35,000 IU of tissue plasminogen activator over 6 h resulted in a significant degree of thrombolysis ($P < 0.001$) whereas infusion of 70,000 IU over 12 h yielded significantly more lysis ($P < 0.01$).

The present findings thus confirm and extend previous *in vitro* observations that tissue plasminogen activator has a much higher fibrinolytic to fibrinogenolytic ratio than urokinase and suggest that it might have greater potential as a thrombolytic agent.

Urokinase was a gift from Lepetit (courtesy of Dr Brioude) and from Choay (courtesy of Dr Toulemonde). D. C. R. was recipient of a fellowship from the Netherlands Organisation for the Advancement of Pure Research (Z.W.O.). This study was supported by grants from the Geconcerteerde Onderzoeksacties (Conventie 80/95-3) and from the ASLK Kankerfonds.

Received 21 November 1980; accepted 1 April 1981.

1. Matsuo, O., Rijken, D. C. & Collen, D. *Thromb. Haem.* (in the press).
2. Rijken, D. C. & Collen, D. *J. biol. Chem.* (in the press).
3. Astrup, T. & Müllertz, S. *Archs Biochem. Biophys.* **40,** 346–351 (1952).
4. Rijken, D. C., Wijngaards, G., Zaal-De Jong, M. & Welbergen, J. *Biochim. biophys. Acta* **580,** 140–153 (1979).
5. Blombäck, B. & Blombäck, M. *Ark. Kemi* **10,** 415–423 (1956).
6. McFarlane, A. S. *Nature* **182,** 53 (1958).
7. Friberger, P. & Knös, M. in *Chromogenic Peptide Substrates* (eds Scully, M. F. & Kakkar, V. V.) 128–140 (Churchill Livingstone, Edinburgh, 1979).
8. Edy, J., De Cock, F. & Collen, D. *Thromb. Res.* **8,** 513–518 (1976).

Demonstration of a specific dynorphin receptor in guinea pig ileum myenteric plexus

Charles Chavkin & Avram Goldstein

Addiction Research Foundation and Department of Pharmacology,
Stanford University, Palo Alto, California 94304, USA

Although the opioid peptide dynorphin contains the pentapeptide Leu-enkephalin as its amino-terminal sequence[1], the two opioid peptides differ in their potency and sensitivity to antagonism by the specific opiate antagonist naloxone. In the guinea pig myenteric plexus–longitudinal muscle assay, dynorphin$_{(1-13)}$ is 700 times more potent and 13 times less sensitive to naloxone antagonism than Leu-enkephalin[1]. We have identified the specific amino acids in the dynorphin$_{(6-13)}$ sequence responsible for this higher potency[2]. One explanation for the difference in naloxone sensitivities is that dynorphin and Leu-enkephalin act through different types of opiate receptors, as first suggested by Goldstein *et al.*[1]. This hypothesis has been strongly supported by recent results[3] showing that vasa deferentia of mice infused with both D-Ala^2-D-Leu^5-enkephalin (DADLE) and the opiate alkaloid sulfentanil become tolerant to the effects of Leu-enkephalin and sulfentanil, whereas sensitivity to dynorphin$_{(1-13)}$ remains unchanged. Here we report that the guinea pig ileum myenteric plexus contains two physically distinct classes of opiate receptors: the μ receptor, with which leu-enkephalin (or normorphine) interacts, and a separate dynorphin receptor.

We used the irreversible antagonist, chlornaltrexamine (CNA), synthesized by Portoghese *et al.*[4] to distinguish the different receptors for dynorphin$_{(1-13)}$ and for Leu-enkephalin. A specific antagonist should be able to selectively block one receptor population. Alternatively, if CNA were not specific, a specific and reversible ligand should be able to protect one receptor population selectively from inactivation by CNA.

CNA acts as a nonequilibrium opiate receptor antagonist[5]. In agreement with this, we have found that in the guinea pig ileum–longitudinal muscle preparation, the extent of antagonism by CNA depends on both its concentration and the duration of exposure. The effects of 20-min exposure to different CNA concentrations on the dose-response curve of Leu-enkephalin are shown in Fig. 1. Low concentrations of CNA (0.5–3.0 nM) caused parallel shifts of the Leu-enkephalin dose–response curve, a result that is not typical of irreversible inhibition[7].

To confirm that CNA is irreversible, 10 muscle strips were washed repeatedly after 20-min exposure to 3 nM CNA. After washing for 60 min, the Leu-enkephalin mean inhibitory concentration IC$_{50}$ was 3.0 μM (16.5-fold increase in IC$_{50}$), and after an additional 4 h washing, the mean IC$_{50}$ was 2.4 μM. The Leu-enkephalin IC$_{50}$ of six muscle strips not exposed to CNA but washed for the same period changed very little (170 nM initially, 270 nM after 5 h). Thus within the time period of the experiment, the antagonism by CNA was irreversible.

Parallel shift of the dose–response curve by an irreversible antagonist indicates that the myenteric plexus contains spare opiate receptors[7] although there are other possible explanations[8]. Creese and Snyder[9] concluded that there are no spare receptors in the myenteric plexus because receptor binding affinity (apparent K_d) was well correlated with pharmacological potency (IC$_{50}$) for a series of non-peptide opiates. Our result might be reconciled with theirs if the parallel shift in the Leu-enkephalin dose–response curve demonstrated here were not due to spare μ receptors (isomorphic spare receptors) but rather due to the presence of receptors of different types (allomorphic spare receptors) in the myenteric plexus.

0028–0836/81/250591—03$01.00

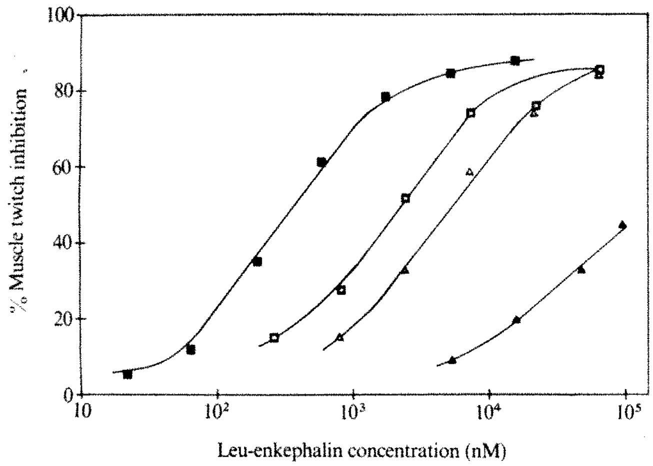


Fig. 1 The effects of different CNA treatments on Leu-enkephalin dose–response curve: ■, before treatment; □, after 0.5 nM CNA; △, after 3 nM CNA; and ▲, after 25 nM CNA. Each point represents the mean per cent inhibition of muscle twitch from 3–4 experiments. Ileum longitudinal muscle from adult male Hartley guinea pigs (Simonsen) was used as described elsewhere[6]. Dilutions of Leu-enkephalin (Biosearch) made in methanol/0.1 M HCL (1:1, v/v) were added to the 5-ml organ bath containing the electrically stimulated muscle strips. A stock solution of 10 μM CNA (given by P. S. Portoghese and A. E. Takemori) in ethanol/0.1 M HCL (1:1, v/v) was stored at −70 °C and fresh dilutions prepared each day. After 20-min CNA exposure, the muscle strips were washed by repeated changes of the Krebs–Ringer buffer at least once every 10 min for 1–2 h before testing Leu-enkephalin. For muscle strips treated with CNA, the concentration of Leu-enkephalin giving 50% inhibition of muscle twitch tension (IC$_{50}$) was increased from 400 nM (untreated strips) to 2.2 μM (0.5 nM CNA), 5.5 μM (3.0 nM CNA), and 173 μM (25 nM CNA).

CNA is a derivative of the reversible antagonist naltrexone, which, like naloxone, is more potent against enkephalin than against dynorphin$_{(1-13)}$ in this tissue. For example we found that the apparent K_d of naltrexone is 3 nM against Leu-enkephalin, but 17 nM against dynorphin$_{(1-13)}$. It was hoped, therefore, that CNA would show a similar differential effect. We exposed the tissue for 20 min to 3 nM CNA then determined the IC$_{50}$ of dynorphin$_{(1-13)}$, Leu-enkephalin and normorphine. As shown in Fig. 2, CNA treatment alone did not discriminate between the receptors occupied by the three ligands. The IC$_{50}$ values were shifted to about the same degree: 14-, 14- and 12-fold, respectively. As was described previously for Leu-enkephalin, the effect of CNA treatment on dynorphin$_{(1-13)}$ IC$_{50}$ was not reversed by repeated washing (4.7 nM, 1 h after CNA; 5.3 nM, 5 h after CNA).

To attempt selective protection, we added either the stable enkephalin analogue, DADLE, or dynorphin$_{(1-13)}$ to the organ bath 1 min before the addition of CNA (3 nM). The lowest concentration of dynorphin$_{(1-13)}$ able to provide substantial protection of dynorphin$_{(1-13)}$ potency was 20 nM, and 100 μM gave complete protection for the 20-min period of exposure to CNA. Comparable concentrations of DADLE were 500 nM and 10 μM for protection of Leu-enkephalin potency. When muscle strips were incubated with the protective ligand (20 nM dynorphin$_{(1-13)}$ or 10 μM DADLE) for 20 min in the absence of CNA and then removed by washing, IC$_{50}$ values were unchanged.

Results of the protection experiments are shown in Fig. 2. When DADLE was present during CNA exposure, the IC$_{50}$ of dynorphin$_{(1-13)}$ was shifted 17-fold—this was not different from the change in IC$_{50}$ in unprotected muscle strips exposed to CNA alone ($P>0.05$, Student's t-test). However, DADLE significantly reduced the IC$_{50}$ ratio of Leu-enkephalin and of normorphine, from 14 to 3 and 12 to 2, respectively ($P<0.01$ for both). Thus the protection afforded by DADLE was selective for Leu-enkephalin and normorphine, compared with dynorphin$_{(1-13)}$. In sharp contrast, when 20 nM dynorphin$_{(1-13)}$ was used as the protecting ligand, the IC$_{50}$ ratio for dynorphin$_{(1-13)}$ was reduced significantly, from 14 to 5 ($P<0.01$), but the IC$_{50}$ ratio for Leu-enkephalin was unaltered. Thus, dynorphin$_{(1-13)}$ provided selective protection of dynorphin$_{(1-13)}$ potency.

These results, showing selective protection of Leu-enkephalin and normorphine potency on the one hand, and dynorphin$_{(1-13)}$ potency on the other, indicate that these ligands act through different binding sites and not through one interconvertible receptor. That DADLE protects both enkephalin and normophine binding sites is consistent with the hypothesis that these

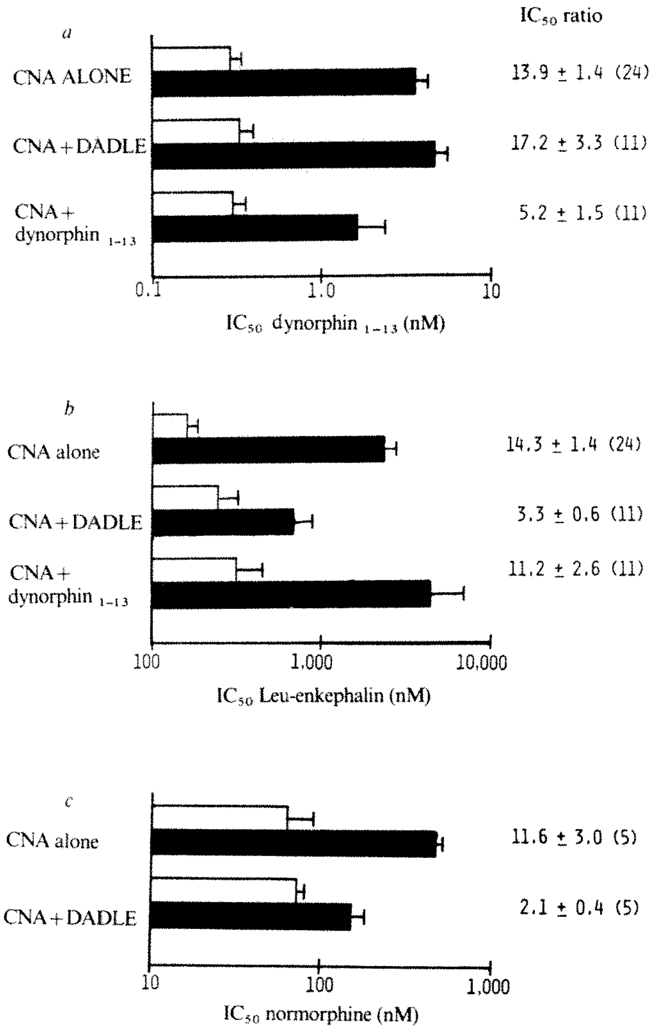


Fig. 2 The effects of CNA treatment on the IC$_{50}$ of *a*, dynorphin$_{(1-13)}$, *b*, Leu-enkephalin, and *c*, normorphine. For each ligand the IC$_{50}$ was estimated before a 20-min incubation of the muscle strip with 3 nM CNA and again after washing for 1–2 h. Open bars represent mean IC$_{50}$ + s.e.m. before CNA treatment; solid bars represent mean IC$_{50}$ + s.e.m. after treatment of 5–24 muscle strips. Dilutions of dynorphin$_{(1-13)}$ (Peninsula), Leu-enkephalin (Biosearch), DADLE (Biosearch), and normophine (Applied Science) were prepared in methanol/0.1 M HCL (1:1, v/v). The IC$_{50}$ was estimated by bracketing the 50% inhibition of the electrically stimulated muscle twitch, and interpolating on a logarithmic concentration–per cent inhibition plot. The protective ligand (either 10 μM DADLE or 20 nM dynorphin$_{(1-13)}$) was added to the bath 1 min before the addition of CNA. Because dynorphin$_{(1-13)}$ is sensitive to peptidases, after 10 min of the 20-min exposure, the strips were washed once and fresh protective ligand and CNA were added. The IC$_{50}$ ratio is the IC$_{50}$ of the agonist after CNA treatment divided by the IC$_{50}$ before CNA treatment as determined for each muscle strip.

two ligands both act through the μ receptor in the guinea pig myenteric plexus, as suggested by Kosterlitz and Paterson[10].

The paradigm of selective protection has been used before to demonstrate differences in opiate receptor types—Robson and Kosterlitz[11] used protection against the alkylating agent phenoxybenzamine and Smith and Simon[12] used protection against *N*-ethylmaleimide to resolve δ and μ binding sites in brain homogenate binding assays. Our study extends that earlier work in two ways. We have used selective protection to distinguish between μ and dynorphin receptors in the guinea pig myenteric plexus and demonstrated that distinction in a functional opiate receptor bioassay system. The presence in guinea pig myenteric plexus of opioid receptors with selective affinity for dynorphin suggests a functional role for this peptide.

We thank Drs P. S. Portoghese and A. E. Takemori for supplying CNA, Drs Brian M. Cox, Francis Leslie, and Carl Romano for helpful discussions, and Pat Lowery for assistance. This work was supported by grants DA-1199 and DA-7063 from the National Institute on Drug Abuse.

Received 27 February; accepted 30 March 1981.

1. Goldstein, A., Tachibana, S., Lowney., L. I., Hunkapiller, M. & Hood, L. *Proc. natn. Acad. Sci. U.S.A.* **76,** 6666–6670 (1979).
2. Chavkin, C. & Goldstein, A. *Soc. Neurosci. Abstr.* **6,** 617 (1980).
3. Wuster, M., Schulz, R. & Herz, A. *Neurosci. Lett.* **20,** 79–83 (1980).
4. Portoghese, P. S., Larson, D. L., Jiang, J. B., Caruso, T. P. & Takemori, A. E. *J. med. Chem* **22,** 168–173 (1979).
5. Caruso, T.P., Larson, D. L., Portoghese, P. S. & Takemori, A. E. *J. Pharmac. exp. Ther.* **213,** 539–544 (1980).
6. Schulz, R. & Goldstein, A. *Br. J. Pharmac.* **48,** 655–666 (1973).
7. Goldstein, A., Aronow, L. & Kalman, S. M. *Principles of Drug Action*, 82–104 (Wiley, New York, 1974).
8. Rocha e Silva, M. *Eur. J. Pharmac.* **6,** 294–302 (1969).
9. Creese, I. & Snyder, S. H. *J. Pharmac. exp. Ther.* **194,** 205–219 (1975).
10. Kosterlitz, H. W. & Paterson, S. J. *Proc. R. Soc.* **B210,** 113–122 (1980).
11. Robson, L. E. & Kosterlitz, H. W. *Proc. R. Soc.* **B205,** 425–432 (1979).
12. Smith, J. R. & Simon, E. J. *Proc. natn. Acad. Sci. U.S.A.* **77,** 281–284 (1980).

Glutamate-preferring receptors regulate the release of D-[^{3}H]aspartate from rat hippocampal slices

Gethin J. McBean & P. J. Roberts*

Department of Physiology and Pharmacology, School of Biochemical
and Physiological Sciences, University of Southampton,
Southampton SO9 3TU, UK

At least two of the excitatory afferent pathways in the hippocampal formation, the Schaffer collaterals which extend from the CA3 to CA1 pyramidal cells and the perforant path from the molecular layer of the dentate gyrus, are believed to use glutamate as their transmitter[1–5]. A third, the commissural pathway, which projects from the CA3/CA4 pyramidals to the contralateral CA1 and dentate areas, may be aspartergic[6]. The excitatory transmitter of the remaining afferent pathway, the mossy fibre projection, is as yet unidentified, but there is no evidence that it is either glutamate or aspartate[6,7]. Despite the obvious importance of excitatory amino acid transmitters in hippocampal function, no evidence of an interaction between these neurotransmitters, such as presynaptic control of release, has been described. Here we report that L-glutamate and several neuroactive analogues inhibit the calcium-dependent, K$^+$-evoked release of D-[^{3}H]aspartate from mini-slices of rat hippocampus. This effect was abolished by selective amino acid antagonists and, as it was tetrodotoxin (TTX) insensitive, the effect may be mediated through presynaptic glutamate autoreceptors.

Superfused hippocampal prisms (0.2 × 0.2 mm), previously loaded with D-[^{3}H]aspartate, attained within 15–20 min a basal release which corresponded to <1% per min of the total accumulated tissue radioactivity (in practice, equivalent to ~3,000 c.p.m. per 10 mg tissue per 5-min perfusion period). The introduction of 50 mM KCl (substituted for sodium) resulted in a >200% increase in the release of the radioactivity. In the presence of a moderately high concentration (100 μM) of L-glutamate, however, the potassium-evoked release of D-[^{3}H]aspartate was markedly attenuated (Table 1), by ~35%. D-glutamate also inhibited the potassium-stimulated release, as did DL-homocysteate, L-homocysteate and L-cysteate. Note that D-homocysteate, which is a strong excitant of central neurones[8], was totally ineffective. L-aspartate, in contrast to glutamate, was without effect, as were *N*-methyl-DL-aspartate, ibotenic acid and its derivative, (*RS*)-α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA). Kainic acid did not affect the release of D-[^{3}H]aspartate. None of the agonists tested exerted any inhibitory effect on the spontaneous release of D-[^{3}H]aspartate. This was not unexpected, as for many other systems regulation of transmitter release by presynaptic receptors can only be demonstrated for the calcium-dependent processes of transmitter release, such as occur with depolarization by either potassium or veratridine. For example, the spontaneous release of ^{3}H-GABA (γ-aminobutyric acid) from rat substantia nigra slices was unaffected by agonists, whereas in the presence of 30 mM K$^+$, a significant inhibitory effect of GABA was observed[9].

One possible complicating factor in the study of presynaptic receptors for amino acid transmitters is that of homoexchange[10], which would result in an apparent enhancement of spontaneous release by agonists which were substrates for the uptake carrier. Although this phenomenon certainly occurs with synaptosomes, it is probably not evident with slices[11] and, certainly in these experiments, enhancement of release was never observed.

Table 1 Effects of excitatory amino acids on the calcium-dependent, K$^+$-evoked release of D-[^{3}H] aspartate from hippocampal slices

Agonist	% Inhibition of K$^+$-stimulated release
L-glutamate (50 μM)	23 ± 18
L-glutamate	35 ± 9*
D-glutamate	32 ± 19†
L-cysteate	41 ± 6‡
L-homocysteate	26 ± 9†
DL-homocysteate	25 ± 13†
(±) Ibotenate	15 ± 10
N-methyl-DL-aspartate	—
Kainate	—
AMPA	—
L-aspartate	—

Rat hippocampal tissue prisms (0.2 × 0.2 mm) were loaded on to gauze rings. One ring was inserted into each of four chambers of a perfusion apparatus[20] and a second ring placed above the first. A tightly fitting nylon piston was pressed down on top of the rings. The tissue was superfused at a constant rate of 0.5 ml min^{-1} with oxygenated Krebs bicarbonate medium (*p*H 7.4, 37 °C containing 0.5 μCi D-[^{3}H]aspartate (specific activity 10 Ci mmol^{-1}) for 15 min to allow high-affinity uptake to occur. Then the slices were rapidly washed and the flow rate increased to 1 ml min^{-1}, and superfusion continued in the absence of label. Released radioactivity was assayed by scintillation counting of 5-min superfusion samples. After achieving a steady basal release rate, 50 mM KCl (substituted for NaCl) was introduced into the perfusion medium for 5 min; this produced a large (213%) calcium-dependent increase in release. L-glutamate (or other analogue) was included with the high-potassium fraction. After stimulation, the medium was exchanged for the low-potassium buffer and a further fraction collected to ensure a return to basal levels. Results are expressed as mean ± s.e.m. of at least four independent determinations. Significance of difference from control K$^+$-evoked release was tested by Student's *t*-test.

$^*P < 0.01$; $^\dagger P < 0.02$; $^\ddagger P < 0.001$.

---

* To whom correspondence should be addressed.

0028–0836/81/250593—02$01.00

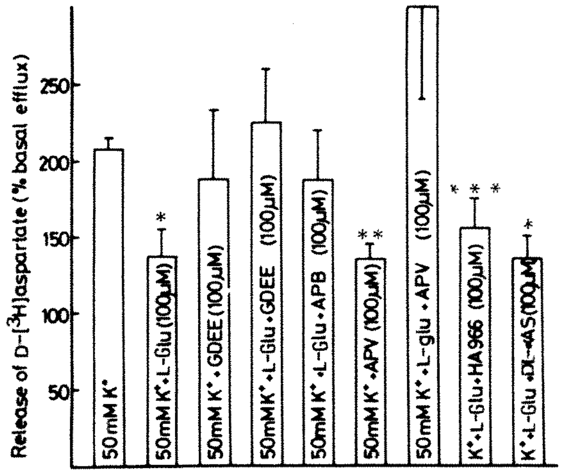


Fig. 1 Effects of amino acid antagonists on the K$^+$-stimulated release of D-[^{3}H]aspartate in the presence of 100 µM L-glutamate. Release of D-[^{3}H]aspartate is expressed as a percentage of basal (unstimulated) efflux. Methods are as described in Table 1 legend. Results are mean ± s.e.m. of at least four determinations. The significance of difference from results obtained with K$^+$ alone was tested by Student's *t*-test. DL-αAS, DL-α-aminosuberate. $*P<0.01$; $**P<0.001$; $***P<0.05$.

The inhibition by 100 µM L-glutamate of the potassium-induced efflux of D-[^{3}H]aspartate could be antagonized by equimolar concentrations of 2-amino-4-phosphonobutyric acid (APB), L-glutamate diethylester (GDEE) and 2-amino-5-phosphonovaleric acid (APV) (Fig. 1). APV alone had a marked glutamate-like action, suggesting that it may act as a partial agonist at these presynaptic receptors. The compounds 1-hydroxy-3-aminopyrrolid-2-one (HA-966) and DL-α-aminosuberate did not affect the inhibition of K$^+$-stimulated D-[^{3}H]aspartate release by L-glutamate, and the response was unchanged in the presence of 0.5 µM TTX in the incubation medium.

These results indicate that there are presynaptic receptors in the hippocampus which modulate the potassium-stimulated release of labelled D-aspartate and show a clear preference for compounds known to interact with glutamate receptors. L-aspartate and *N*-methyl aspartate (an aspartate receptor-preferring agonist[12]) were not active. (±) Ibotenate, which may also interact preferentially with aspartate receptors[13], was also ineffective in this system. Kainic acid, although it is a potent neuroexcitatory analogue of glutamate, does not seem to interact directly with postsynaptic glutamate receptors[14]. Similarly, in this system, there is no evidence for an interaction between kainate and the presynaptic glutamate receptor. With regard to the use of antagonists, it is also significant that HA-966 and DL-α-aminosuberate were inactive as both show marked selectivity for aspartate receptors[15], compared with GDEE, for example, which is strongly glutamate receptor-preferring.

Although D-[^{3}H]aspartate is a valuable tool for studying excitatory amino acid release mechanisms because it is taken up into both glutamatergic and aspartergic terminals as effectively as the natural L-isomers, it is essentially not metabolized by nervous tissue, including the hippocampus[16–18]. Thus, although we have identified a presynaptic receptor for glutamate, we cannot say whether this is an autoreceptor or a glutamate receptor which regulates the release of L-aspartate. We have recently begun to investigate the potassium-evoked, Ca^{2+}-dependent release of endogenous L-glutamate from hippocampal slices[19] and have found that transmitter release is again significantly decreased in the presence of glutamate analogues. This implies that, at least in the hippocampus, glutamate regulates its own release through autoreceptors. However, we have been unable to detect a similar system in either striatum or cerebellum.

This work was supported by a SRC grant to P.J.R. We thank Drs J. C. Watkins and P. Krogsgaard-Larsen for gifts of compounds.

Received 21 January; accepted 3 April 1981.

1. Malthe-Sørenssen, D., Skrede, K. K. & Fonnum, F. *Neuroscience* **4,** 1255–1263 (1979).
2. Nadler, J. V., White, W. F., Vacca, K., Redburn, D. A. & Cotman, C. W. *J. Neurochem.* **29,** 279–290 (1977).
3. Hamberger, A., Chiang, G., Nylen, E. S., Scheff, S. W. & Cotman, C. W. *Brain Res.* **143,** 549–555 (1978).
4. Hamberger, A., Chiang, G., Nylen, E. S., Scheff, S. W. & Cotman, C. W. *Brain Res.* **148,** 513–530 (1979).
5. Hamberger, A., Chiang, G., Nylen, E. S., Scheff, S. W. & Cotman, C. W. *Brain Res.* **168,** 531–541 (1979).
6. Nadler, J. V., White, W. F., Vacca, K., Perry, B. W. & Cotman, C. W. *J. Neurochem.* **31,** 147–155 (1978).
7. White, W. F., Nadler, J. V. & Cotman, C. W. *Brain Res.* **164,** 177–184 (1979).
8. Balcar, C. J. & Johnston, G. A. R. *J. Neurobiol.* **3,** 295–301 (1972).
9. Kamal, L., Arbilla, S. & Langer, S. Z. in *Presynaptic Receptors* (eds Langer, S. Z., Starke, K. & Dubocovich, M. L.) 193–197 (Pergamon, Oxford, 1979).
10. Levi, M. & Raiteri, M. *Int. Rev. Neurobiol.* **19,** 51–74 (1976).
11. Johnston, G. A. R. in *Glutamate, Transmitter in the Central Nervous System* (eds Roberts, P. J., Storm-Mathisen, J. & Johnston, G. A. R.) 77–87 (Wiley, Chichester, in the press).
12. McCulloch, R. M., Johnston, G. A. R., Game, C. J. A. & Curtis, D. R. *Expl Brain Res.* **21,** 515–518 (1974).
13. McLennan, J. & Lodge, D. *Brain Res.* **169,** 83–90 (1979).
14. Foster, A. C. & Roberts, P. J. *J. Neurochem.* **31,** 1467–1477 (1978).
15. Davies, J. & Watkins, J. C. in *Glutamate as a Neurotransmitter* (eds Di Chiara, G. & Gessa, G. L.) 275–284 (Raven, New York, 1980).
16. Davies, L. P. & Johnston, G. A. R. *J. Neurochem.* **26,** 1007–1010 (1976).
17. Lund Karlsen, R. & Fonnum, F. *Brain Res.* **151,** 457–467 (1978).
18. Storm-Mathisen, J. & Woxen Opsahl, M. *Neurosci. Lett.* **9,** 65–70 (1978).
19. McBean, G. J. & Roberts, P. J. (in preparation).
20. Rowlands, G. J. & Roberts, P. J. *Expl Brain Res.* **39,** 239–240 (1980).

Transcription of immunoglobulin heavy-chain sequences from the excluded allele

**Phyllis A. Ponte*, Miriam Siekevitz†,
Richard C. Schwartz†,
Malcolm L. Gefter† & Gail E. Sonenshein***

*Department of Biochemistry, Boston University School of Medicine, Boston, Massachusetts 02118, USA
†Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Production of a functional immunoglobulin heavy-chain gene requires multiple recombination events[1–5]. In germ-line cells, segments of the classical variable portion are encoded in separate genetic elements: variable (V), diversity (D) and joining (J) regions[3,5,6]. The genes encoding the different clustered constant-region domains (5′-µ, δ, γ^3, γ^1, γ^{2b}, γ^{2a}, ε, α-3′) are widely separated from the variable region[7–12]. During differentiation to immunoglobulin production, one allele containing genes coding for the V, D, J and C segments is rearranged to form a coding sequence for the synthesis of heavy chain. In some cells, the other alleles for heavy-chain segments are rearranged, albeit differently[8]. At the protein level, expression of only one of the alleles is detected (allelic exclusion)[13,14]. We report here studies on a class of mouse myeloma MOPC 315 variant cell lines which has lost the ability to produce α heavy-chain protein or mRNA. Hybridization analysis indicates that the functional α allele has been lost. However, these variants synthesize unusual RNA species that include α constant-region sequences. We conclude that this transcription is from the 'excluded' allele. Expression of these aberrant RNAs is altered in the presence of a productive gene.

Wild-type mouse myeloma MOPC 315 cells synthesize an immunoglobulin A (IgA, λ^2) protein. We have isolated several variant lines which have lost the ability to produce α heavy-chain protein but have normal levels of λ^2 light chain. To test for the presence of heavy-chain mRNA, cytoplasmic RNA was

0028-0836/81/250594—03$01.00

translated in a cell-free system and the immunoprecipitated products separated by electrophoresis. RNA from the wild-type cell line (LK) directs the synthesis of both heavy- and light-chain polypeptides (Fig. 1). RNAs from all of the variant lines tested (which includes three additional lines not shown) are unable to direct heavy-chain protein synthesis but normal levels of light chain are produced. Heavy-chain translation at about 1/10th the normal level would have been detected.

To test for the presence of expressed α constant-region sequences, poly(A)-containing cytoplasmic RNA, bound to diazobenzyloxymethyl (DBM)-paper[15], was hybridized to ^{32}P-labelled cloned DNA. With the α probe, RNA from wild-type LK cells shows a strong band of hybridization at ~1,950 bases (Fig. 2*a*, *c*). This is the appropriate size for a full-length α heavy-chain mRNA. However, in addition to this mRNA band, there are low levels of smaller molecular weight RNAs of discrete size. To determine whether these RNA species resulted from degradation due to experimental manipulation, the α DNA probe was removed from the DBM filter[15] and the filter rehybridized to λ^2 DNA probe. LK and all the other lines tested show a single strong hybridization band at 1,200 bases, the size of λ^2 mRNA (Fig. 2*b*). Surprisingly, when RNA extracted from five of the independently isolated heavy-chain non-producing variant lines is hybridized to α DNA probe, a complex series of discrete bands is seen (Fig. 2*a*), none of which seem to co-migrate with α heavy-chain mRNA. The lower bands (1 and 2) apparently correspond to species seen in wild-type cells. Four additional RNA species are also detected; one of these (3) has only a slightly greater mobility than full-length RNA. The other three bands are over full-length size; one is only slightly larger (4) and two are high molecular weight species (5 and 6 are 5,700 and 6,900 bases, respectively). As no nuclear precursor species are detected on hybridization with λ^2 probe (Fig. 2*b*), it is unlikely that the α RNAs have leaked into the cytoplasm during isolation.

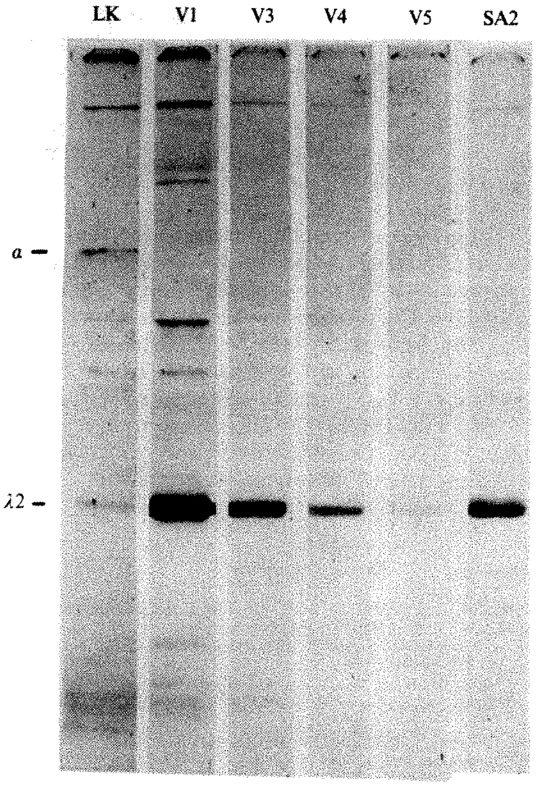


Fig. 1 Cell-free translation analysis of wild-type and variant cell line RNA. MOPC 315 heavy-chain non-producing variant cell lines were isolated from LK (wild type) cells which had been treated with the following mutagens: V1 and V4, ICR 191-OH; V2 and V3, ICR 191-F; V5 was first treated with ethylmethane sulphonate (EMS), and then with ICR 191-F. After mutagenesis, cells were cloned in agar and variant lines isolated using specific antisera to MOPC 315 IgA protein in an overlay assay[25]. Radioimmunoassays, performed as described elsewhere[16], failed to detect any full-length α heavy-chain protein intracellularly. The assay can detect 10^{-4} times the amount of heavy chain present in steady-state wild-type cells. Cell line SA2 was previously isolated by the same overlay procedure as a spontaneously arising variant of a hybrid cell line derived from fusion of MOPC 315 with MPC 11. The hybrid line (SOO-1) produces all four parent immunoglobulin chains, α, λ^2, γ^{2b} and κ, whereas SA2 has lost α-chain production but continues to produce the other chains[16]. Cytoplasmic RNA was extracted and translated in a rabbit reticulocyte cell free system[26,27]. The products were precipitated with rabbit anti-315 protein followed by goat anti-rabbit immunoglobulin antisera[26]. Samples were prepared and electrophoresed on 5–15% SDS-polyacrylamide gels[26], which were treated for fluorography[28].

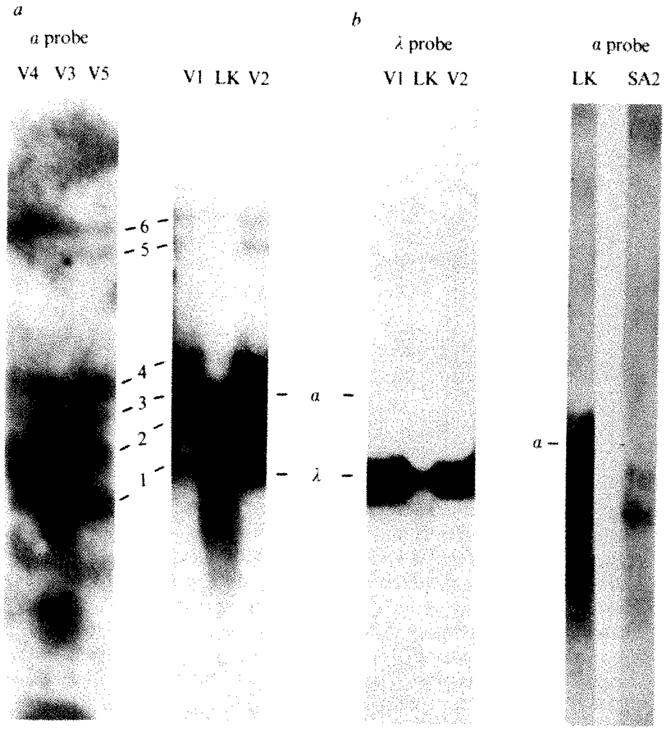


Fig. 2 Comparison of RNA species from wild-type and variant cells. Poly(A)$^+$-selected cytoplasmic RNA was denatured with glyoxal[29] and electrophoresed on 1.1% agrose gels. The tracks containing LK (wild type) RNA had 2 μg of RNA, and all those containing variant RNA had 4 μg. RNA transfer to DBM-paper and hybridization procedures were as described elsewhere[15]. The ^{32}P-labelled probes used were prepared by nick-translation[30] to specific activities of $1\text{–}2 \times 10^8$ c.p.m. per ug and $\sim 5 \times 10^6$ c.p.m. were hybridized per track. Position of wild-type full-length heavy- and light-chain mRNAs are indicated as α- and λ-. Visual markers used were rRNA and its precursors prepared from myeloma nuclei. The cloned α heavy-chain DNA used these experiments is 650 base pairs long and corresponds to the 3′ end of the heavy-chain mRNA as far as the middle of the constant region. The 870-base pair λ^2 insert is complementary to 3′-end sequences and extends well into the variable region of the light chain. Plasmid pBR322 clones containing inserts of cDNA to MOPC 315 heavy-and light chain mRNA were constructed as described previously[20]. Cloned DNAs were characterized by hybridization to specific mRNA ('positive selection')[31] and by hybridization arrest of translation (HART) analysis[32]

To determine whether the production of aberrant RNA is reflected in the heavy-chain gene organization, we compared the restriction patterns for α-chain coding sequences in wild-type and variant lines. Restriction of wild-type MOPC 315 DNA with *Bam*HI produces two restriction fragments of 12.5 and 9.3 kilobases, containing α constant-region sequences (Fig. 3*a*); the germ-line restriction pattern gave a single band at 16.5 kilobases[8]. DNA from the heavy-chain non-producing mutant cells

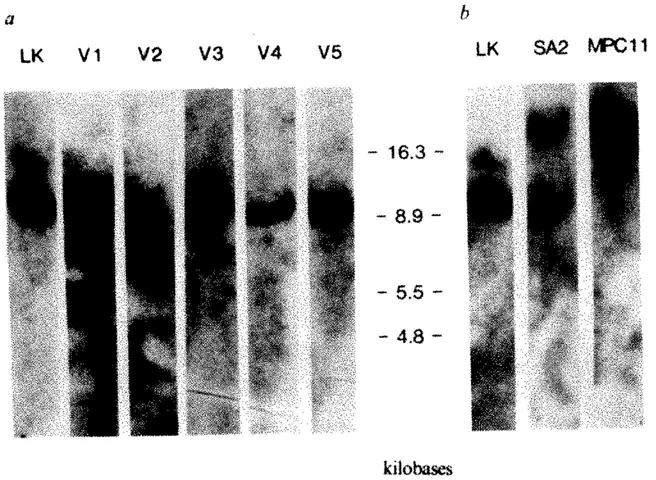


Fig. 3 Hybridization of *Bam*HI-digested DNA to α constant-region sequences. DNA was purified from cells and restricted to completion with *Bam*HI endonuclease. After electrophoresis on a 0.9% agarose gel, the DNA (10 μg per slot) was transferred to nitrocellulose filters (Millipore) according to Southern[33]. Hybridization mixtures contained 5×10^5 c.p.m. per slot of a ^{32}P-labelled α constant-region probe (2×10^8 c.p.m. per ug); hybridization was as described elsewhere[34]. Washes were at 68 °C with 3×SSC, 0.2% SDS, 10×Denhardts for 1 h, 2×SSC, 0.2% SDS for 1 h followed by 3×1 h with 1×SSC, 0.2% SDS. (1×SSC is 0.15 M NaCl, 0.015 M Na-citrate; 1×Denhardts is 0.02% each of Ficoll 400, polyvinyl pyrrolidone and bovine albumin.) Fragment sizes were determined by comparison with λ phage DNA restriction fragments.

have lost the 12.5-kilobase restriction fragment (Fig. 3*a*). This result is consistent with the 12.5-kilobase band containing the gene coding for the production of heavy-chain mRNA and the 9.3-kilobase band containing the 'allelically excluded' gene. Hybridization with J and switch-region probes has confirmed that the 12.5-kilobase band is the functional allele[35]. Therefore the aberrant α RNA species in the variants are likely to be due to transcription of the rearranged, but allelically excluded, heavy-chain gene.

Cory and Adams[8] have shown that the DNA around the α constant region has rearranged in some myelomas producing heavy chains which map upstream to the α constant region, for example, γ^{2b} and γ^1. Is such a rearrangement of an α heavy chain gene sufficient to promote transcription? MPC 11 (IgG^{2b}), in which one α constant-region allele is rearranged from the germ-line configuration[8], was tested for the presence of α-specific RNA sequences. No detectable levels of material hybridizing to α constant-region DNA were found in cytoplasmic poly(A)$^+$ RNA preparations from this line (data not shown), nor could we detect any transcription of α constant-region sequences in MOPC 21 (IgG1), whose α constant-region genes both remain in the germ-line configuration[8].

To determine whether the presence of a functional heavy-chain gene other than that for α heavy chain would influence the expression of the excluded allele, RNA from a hybrid line was examined. SA2 is a variant line, isolated from an MPC 11× MOPC 315 hybrid cell[16], which has lost the ability to produce α heavy chain (Fig. 1). It synthesizes γ^{2b} heavy chain and a κ light chain from the MPC11 parent and a λ^2 light chain from the MOPC 315 parent.

Cytoplasmic RNA from SA2 contains low levels of the low-molecular weight α constant-region species (Fig. 2*c*). Interestingly, the RNA species that are larger than full-length message (species 4, 5 and 6) are not detected. Southern hybridization analysis reveals a band at ~16.5 kilobases derived from the MPC 11 parent[8] and a band at 9.3 kilobases from the MOPC 315 parent (Fig. 3*b*). The 12.5-kilobase productive α allele is absent. Thus transcription of the 9.3-kilobase allele occurs in the SA2 line. The presence of an active γ^{2b} gene does not 'turn off' this activity, although the levels of the aberrant RNAs detected are reduced.

Transcription of light-chain gene sequences to give species other than the productive message has recently been described in myeloma cells[17-19]. MOPC 315 cells produce a fragment of λ^1 mRNA, as well as full-length λ^2 mRNA[20,21]. Our results indicate that wild-type MOPC 315 cells synthesize, in addition to a full-length heavy-chain mRNA, lower molecular weight RNA species which contain α constant-region sequences. Similar RNAs are found in variant lines that are unable to produce full-length mRNA due to the absence of the productive allele. Curiously, RNA species of a higher molecular weight than full-length heavy-chain mRNA are also detected in mutant lines. We propose that these species represent transcription from the allelically excluded gene but the nature of these transcripts and their function in the cell is unknown.

Note that the species of greater than full-length message size are only found in cells which produce no heavy-chain protein. They are not detected in the wild-type MOPC 315 or in two other mutant cell lines with altered heavy-chain production (one of which overproduces heavy-chain protein and the other an aberrant heavy chain (manuscript in preparation)). They are also not detected in the hybrid line, SA2, which produces a functional γ^{2b} chain. Although it is not understood how they arise, it may be that the steady-state levels of these aberrant RNAs are regulated by the presence of a functional heavy-chain gene or gene product. This regulation could occur at the level of transcription or RNA splicing. Experiments are being done to determing if these RNA species contain the J region or unspliced α-chain intervening regions[22-24]. Finally, note that in this case of heavy-chain RNA, and in all the reports of aberrant light-chain transcription[17-21], allelic exclusion at the level of productive chain synthesis is always maintained.

We thank Dr S. Cory for sharing her results before publication and James McCormack for technical assistance. G.E.S. is an established investigator of the American Heart Association. P.A.P., R.C.S. and M. S. were supported by NIH training grants T32HL07052 and GM 07287. This research was supported by NIH grant AI16051 and American Cancer Society grant IM-214. All experiments involving recombinant DNA were conducted in accordance with NIH guidelines.

Received 9 December 1980; accepted 25 March 1981.

1. Dreyer, W. J. & Bennett, J. C. *Proc. natn. Acad. Sci. U.S.A.* **54,** 864–869 (1965).
2. Sakano, H. *et al. Nature* **277,** 627–633 (1979).
3. Early, P. W., Davis, M. M., Kaback, D. B., Davidson, N. & Hood, L. *Proc. natn. Acad. Sci. U.S.A.* **76,** 857–861 (1979).
4. Early, P., Huang, H., Davis, M., Calame, K. & Hood, L. *Cell* **19,** 981–992 (1980).
5. Sakano, H., Maki, R., Kurosawa, Y., Roeder, W. & Tonegawa, S. *Nature* **286,** 676–683 (1980).
6. Schilling, J., Clevinger, B., Davie, J. M. & Hood, L. *Nature* **283,** 35–40 (1980).
7. Honjo, T. & Kataoka, T. *Proc. natn. Acad. Sci. U.S.A.* **75,** 2140–2144 (1978).
8. Cory, S. & Adams, J. M. *Cell* **19,** 37–51 (1980).
9. Cory, S., Jackson, J. & Adams, J. M. *Nature* **285,** 450–456 (1980).
10. Shimizu, A. *et al Nature* **289,** 149–153 (1981).
11. Rabbitts, T. H., Forster, A., Dunnick, W. & Bentley, D. L. *Nature* **283,** 351–356 (1980).
12. Yaoita, Y. & Honjo, T. *Nature* **286,** 850–853 (1980).
13. Pernis, B., Chiappino, G., Kelus, A. S. & Gell, P. *J. exp. Med.* **122,** 853–875 (1965).
14. Cebra, J. S., Colberg, J. E. & Dray, S. *J. exp. Med.* **123,** 547–558 (1966).
15. Alwine, J. C., Kemp, D. J. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5350–5354 (1977).
16. Siebert, G. R., Harris, J. F. & Gefter, M. L. *J. Immun.* **121,** 1808–1818 (1978).
17. Seidman, J. G. & Leder, P. *Nature* **276,** 790–795 (1978).
18. Perry, R. P. *et al. Proc. natn. Acad. Sci. U.S.A.* **77,** 1937–1941 (1980).
19. Alt, F. W., Enea, V., Bothwell, A.L.M. & Baltimore, D. *Cell* **21,** 1–12 (1980).
20. Bothwell, A. *et al. Nature* **290,** 65–67 (1981).
21. Schwartz, R. C., Sonenshein, G. E., Bothwell, A. & Gefter, M. L. *J. Immun.* (in the press).
22. Davis, M. M. *et al. Nature* **283,** 733–739 (1980).
23. Gilmore-Hebert, M. & Wall, R. *Proc. natn Acad. Sci. U.S.A.* **75,** 342–345 (1978).
24. Schibler, U., Marcu, K. B. & Perry, R. P. *Cell* **15,** 1495–1509 (1978).
25. Coffino, P. & Scharff, M. *Proc. natn. Acad. Sci. U.S.A.* **68,** 219–223 (1971).
26. Sonenshein, G. E., Siekevitz, M., Siebert, G. R. & Gefter, M. *J. exp. Med.* **148,** 301–312 (1978).
27. Pelham, H. R. & Jackson, R. J. *Eur. J. Biochem.* **67,** 247–256 (1976).
28. Bonner, W. M. & Laskey, K. A. *Eur. J. Biochem.* **46,** 83–88 (1974).
29. McMaster, G. K. & Carmichael, G. C. *Proc. natn. Acad. Sci. U.S.A.* **74,** 4835–4838 (1977).
30. Rigby, P. W. J., Diekmann, M., Rhodes, C. & Berg, P. *J. molec. Biol.* **113,** 237–251 (1977).
31. Ricciardi, R., Miller, J. & Roberts, B. *Proc. natn. Acad. Sci. U.S.A.* **76,** 4927–4931 (1979).
32. Paterson, B. M., Roberts, B. E. & Kuff, E.L. *Proc. natn. Acad. Sci. U.S.A.* **74,** 4370–4374 (1977).
33. Southern, E. M. *J. molec. Biol.* **98,** 503–517 (1975).
34. Jeffreys, A. J. & Flavell, R. A. *Cell* **12,** 429–439 (1977).
35. Cory, S., Webb, J. & Adams, J. *Biochemistry* (in the press).

BOOK REVIEWS

Why do as you would be done by?

Vernon Reynolds

The Expanding Circle: Ethics and Sociobiology. By Peter Singer. Pp.190. ISBN 0-374-15112-1. (Farrar, Straus & Giroux: 1981.) $10.95. To be published in Europe and the Commonwealth on July 30th by Oxford University Press, £6.95.

MERE mention of the word "ethics" tends to make people's eyes go glassy. Yet, well trodden as it is, the field of ethical philosophy is wide open to new ideas. Indeed, Singer believes we are still only at the "beginning of a long march", only now obtaining the "knowledge [that] may clear up the confusion that surrounds ethics" (p.168).

The new knowledge that Singer injects into ethical philosophy in this book is the theory and related facts of biological altruism. Without defining ethics as such, he treats it for the most part (except in a section towards the end) as a quality of personal action, that quality which is concerned not so much with what we do, or why we do it, as with our evaluation of it, whether it is right or wrong, good or bad. As he shows, with great perseverance and lucidity, this evaluation is independent of the actions themselves, and therefore no amount of objective demonstration of the causes of our acting in particular ways can explain ethics.

But Singer is aware that this tends to leave ethics "hanging in the air" as a rootless *a priori,* and this, he feels, is a weakness. In what, if anything, might ethics be rooted? In religion? In social rules? In rational thought? In biology? All have a share, but the deepest roots are in biology. The theory of kin and reciprocal altruism, while wholly unable to give form or content to ethical evaluations, can, he believes, help us to understand the human being as an ethical animal, by showing that many species have, for genetic reasons, evolved patterns of behaviour which foster the survival and reproductive success of others in their own kin group or social group. To this extent "ethics is part of the natural human condition" (p.23).

Singer denies that there is any general difference between human beings and animals. We are all part of a common evolutionary process. We all share "common elements" (p.29) and these have a biological basis. What these elements are is not specified, but genes are frequently referred to, and Singer is clear that mankind cannot escape from the process of natural selection. He is equally clear, though, that in the case of such a complex information-processing and often creatively thinking organism as *Homo sapiens,* the appeal to genes as such for the explanation of right–wrong evaluations is falsely based. What then is the relation between our ethics and biological altruism?

First we should look at Singer's biology. He outlines his theory of altruism in Chapter 1. It is gene-based, and adds up the combined positive effects of an organism's behaviour on the survival-to-reproductive-age chances of its kin and neighbours to establish the probability of transmission. Here he goes, rightly, beyond the Lorenzian concern with motivation-structure (aggression, fear, sex) to find the roots of altruism not in the outcome of motivated decisions about behaviour but in the inevitable results of the genetic process, the blind "logic of living systems" as François Jacob calls it. Wilson, Dawkins, Trivers, Hamilton, Smith, Prince, Parker, all have their share in the ideas expressed.

Of these, Wilson gets the lion's share of attention. He comes under fire from Singer in his claim that a biology of ethics will do away with the need for an independent ethics (and ethical philosophers), being able to settle matters of ethics by reference to predictable biological outcomes. This, for Singer, overstates the place of biology in relation to ethics. Values come first, and then the predicted outcomes of actions and all other relevant knowledge can be assessed, and plans made and acted on. Singer does not accept the use of scientific theories or data as a basis for an ethics of the here-and-now; this for him is a case of the naturalistic fallacy, moving from "is" to "ought", a logically indefensible step first shown to be such by Hume in 1739. Whereas science can provide explanations and facts, ethics consists of directives. Facts have no power to direct; they are neutral; ethics has power to direct but cannot provide facts or explanations. We need ethics and facts for rational planning at all levels — personal, group, race or species.

The matter of levels clearly concerns Singer, perhaps because of the emphasis in evolutionary biology on the individual or the gene as the normal unit of selection. In striving for a unified theory that will link here-and-now ethics to its biological base, there is not just the problem of priorities discussed above, but the problem of how in a process based on differentials between individuals there can arise an ethics transcending individual interests and reaching to those of the group, the species and even — remember that Singer was author of *Animal Liberation* (Cape, 1976) — of other species. This is the problem of the expanding circle: how, he wants to know, has ethics come, in many cultures and at various times, to transcend itself, from personal to tribal considerations, from tribal to racial, from racial to mankind, from mankind to all sentient beings? Are all such extensions outgrowths of the sentiments evolved as a result of kin and reciprocal altruism? And if so, what is the mechanism of the process?

Singer argues that the mechanism is the application of rational thought and action, which can transcend the situations provided by the kin-group and co-residential group. The process is the growth of universal values out of small-group values. The pay-offs, though not specified by Singer, are presumably pay-offs at the whole-group, national or even wider levels, in terms of political stability and economic security. The rational element he stresses is the exhortation "do as you would be done by". On this basis, people act not only for themselves, or their kin. They also act for their groups, their nationals everywhere, the members of other races, all members of the species and even members of other species involved in the common life-process on our planet (other planets aren't considered!). All can be brought under a set of ethical principles; indeed, he sees this as "progressive" (p.117).

But because of the expansion of ethics in this way we should not suppose that one or two clear-cut directives will suffice; that would be abandoning rational ethics for an ethics based on intuition or ancestral custom or some other principle. Singer is not in favour of such a rigid ethics; he wants a flexible, situation-specific ethics that can take account of the complexity of the world. Just as Mary Midgley, seaching for the common good, concluded that *bonum est multiplex,* Singer concludes that the rules of ethics are at all times "subject to critical scrutiny" (p.164).

The book is a short one. Because of the felicitous style, it is an easy read. But it is packed with ideas needing further exploration.

In particular, the ethics of "do as you would be done by" has to be squared with the existence of groups who would be done by in different ways. One is reminded, at the individual level, of the dilemma faced by the young queer from Khartoum, who took a lesbian up to his room. . . . On the wider level, do we do unto Nazis, knowing their ideological bent, as we would be done by? Singer, to be sure, wants flexibility. But how much, when and where? Has he really validated his title? Are there not times when the circle must be contracted, to nation, family or self? I think Singer will readily agree that there are, but he has not pursued this aspect as fully as it deserves; one could start, perhaps, from the work of Coser on the functions of social conflict, or Evans Pritchard's analysis of segmentary social systems.

Second, anthropologists must surely find it hard to accept the approach to "rationality" adopted by Singer. He speaks of rationality as if it were a unitary phenomenon, leading properly, given a particular set of considerations, to certain conclusions. Not only is this all too rarely the case within cultures sharing one set of concepts — one language etc. — but between cultures the idea of what it is to think or behave rationally varies in many dimensions. It is doubtful indeed whether all cultures even have a notion of rationality. Who knows but it may be advantageous, in order to overcome the "cognitive dissonance" problem he refers to (p.143), to dispense with it completely?

Criticisms aside, however, this book is undoubtedly an excellent exposition of a most complex problem. Peter Singer has without question clarified the many subtleties and singularities involved in the relation between ethics and sociobiology. ☐

Vernon Reynolds is University Lecturer in Physical Anthropology at the University of Oxford, and author of The Biology of Human Action (*W.H. Freeman, 2nd Edn, 1980*).

Imagination more firmly controlled

Morris Moscovitch

Mental Imagery and Human Memory. By J.T.E. Richardson. Pp.178. ISBN 0-333-25370-1/0-312-52975-9. (Macmillan, London/St Martin's Press: 1980.) £16, $22.50.

AFTER being neglected for about a half-century by experimental psychologists in the English-speaking world, mental imagery once again become a worthy subject of inquiry in the middle 1960s. The fact that such private, mental representations could be studied objectively signalled the end of the old stimulus–response behaviourism and augered the birth of a new cognitive psychology that could explore freely a variety of phenomena that are part of mental life.

By the mid-1970s, however, the new cognitive psychology had become the established doctrine and research on mental imagery had grown duller while, at the same time, the controversy surrounding it had become more intense. Currently, much research on mental imagery seems to be concerned with establishing the nature of the code in which mental images are represented rather than in exploring further the nature of the phenomenon or the uses to which it can be put. One group believes that mental images are represented in a code that is fundamentally different from that in which verbal information is represented. The former is concrete, analogical and perhaps isomorphic with perception, whereas the latter is abstract and propositional. Against this dual-coding position is the view that all information is represented in terms of a common code involving networks of propositions. The controversy has taken on the trappings of an arms race. No sooner has one side unveiled a powerful empirical or theoretical weapon which threatens the other with ultimate defeat, than an equally destructive finding or argument is developed by the other side. One can only lament this expenditure of talent, because the differences between the two camps, though large in theory, have become almost indistinguishable in practice. It often takes more than an expert to decide whether an experiment struck a blow for one side or the other.

Such expert refereeing is provided by J. T. E. Richardson in *Mental Imagery and Human Memory.* As the title suggests, he is concerned with mental imagery only insofar as it contributes to human memory. His aims, though limited, are not modest. They are to analyse and comment on conceptual and methodological problems and to integrate and organize the research findings in an area that has grown exponentially over the years and has become quite unwieldy as a result. By and large he has done a good job.

In the early chapters he reviews critically, yet succinctly, the major historical trends in imagery research that led to the current state of affairs. He suggests that some of the conceptual problems facing imagery research arose in part as a result of making imagery acceptable to a behaviouristically orientated psychology. By providing a conceptual analysis of mental imagery as it is used in ordinary language and psychological discourse, Richardson shows up the shortcomings of some of the current approaches and clarifies the theoretical arguments between the single- and dual-coding theorists. Particularly important here is the idea that mental images, because they are representations, have the logical feature of intentionality. That is, the object and functional origin of the image is determined by how it is intended, not by what it resembles. Thus, images arise from "a system of knowledge which is conceptual and propositional" (p.40). Once constructed, however, images may have emergent properties which could not be computed from the original descriptions and which may not be coded in propositional form.

Thus armed, Richardson reviews the experimental literature on imagery and memory. Chapters 4 and 5 are on imagery and immediate memory and on pictorial imagery, Chapters 6–8 deal primarily with imagery and long-term memory, and the last chapter concerns individual differences. He argues that whereas the effects of imagery on long-term memory for single words and sentences are best explained by a propositional theory, the use of imagery in immediate memory (such as rotating or inspecting a figure mentally) and in remembering pictures is best explained in terms of a dual-coding theory. This suggests that mental images may have different functional properties in different instances and, consequently, the evidence from these instances should be weighed separately in deciding theoretical issues.

I regret that he chose not to discuss experiments in the phenomenology of imagery (how big is an image, how wide an angle does it subtend, how long does it take to conjure it up?), since it would have fitted nicely into the chapter on imagery and working memory. I also wish that he had dealt more fully with neuropsychological approaches to problems in mental imagery. Recent investigations by Bisiach and Luzzatti (*Cortex* **14**, 129–133; 1978), who found that individuals with left-sided neglect also neglect the left side of their mental image, suggest that such investigations have a useful contribution to make.

However, the book is an excellent and very readable source of information on research in mental imagery. Richardson has acted as referee and thereby clarified and made more interesting an area that was badly in need of organization. It is left to another book to offer new directions to research and theory ☐

Morris Moscovitch is an Associate Professor in the Psychology Department and Centre for Research in Human Development at Erindale College, University of Toronto.

Estuarine mixture

J. D. Burton

Chemistry and Biogeochemistry of Estuaries. Edited by E. Olausson and I. Cato. Pp.452. ISBN 0-471-27679-0. (Wiley-Interscience: 1980.) £28.50, $78.40.

THE study of estuaries presents a notable challenge to aquatic geochemists because of the range of physicochemical conditions and variations in composition encountered in estuarine waters and sediments, and the variability and complexity of estuarine environments. There have been several reasons for taking up this challenge and for the consequent emergence, during the past decade, of estuarine chemistry as a distinctive field. The complex pathways of chemical species in estuaries provide important examples of the operation of many geochemical processes. Knowledge of the source and sink functions of estuaries is crucial to an understanding of the chemical mass balances in the shelf seas and oceans. In addition to these fundamental aspects, the extensive — and sometimes conflicting — uses of many estuaries for industrial, commercial and recreational purposes, have led to a need, in terms of resource management, to understand the estuarine behaviour of pollutants.

The need for multidisciplinary awareness in the study of environmental chemical processes is particularly apparent in estuarine chemistry, and the present book appropriately contains four chapters which deal with geomorphology and classification of estuaries, physical processes of water transport and mixing, and sedimentary and biological processes. They serve very well the purpose of providing background, and will also be consulted in their own right. The wisdom of treating the chemistry in as many as nine further chapters may be questioned. Despite the degree of subdivision — which means that the emerging coherence of the subject tends to be obscured — there remain significant omissions and imbalances which are aggravated by the problem, common to many edited works, of uneven treatment. Thus, although there are two contributions dealing with specific groups of organic pollutants (halogenated hydrocarbons and petroleum hydrocarbons), the more general organic chemistry of estuaries, including the role of humic materials, is largely overlooked. The use of radionuclides in the natural decay series to study estuarine processes is scarcely discussed, certainly not commensurately with the importance of the insights which have been obtained in this way. The carbon dioxide system, a key topic, is accorded a separate chapter but nevertheless receives remarkably cursory treatment.

The strengths lie mainly in the greater than usual account taken of the extensive European literature on the subject, and in the high quality of some individual chapters. Among these, there is a lucid treatment of equilibrium modelling of chemical speciation by Dyrssen and Wedborg, a well integrated and well documented discussion by Duinker of suspended particles and their interactions, and a welcome account by Presley and Trefry of intersititial waters of sediments. An unusual and interesting topic is the recovery and decontamination of polluted estuaries, discussed by Cato, Olsson and Rosenberg, with particular emphasis on benthic communities. The strong points are sufficient to make the book a useful addition to the literature of estuarine science and one which should interest a wider audience than just the specialists in estuarine chemistry. □

J.D. Burton is Reader in Chemical Oceanography at the University of Southampton.

Hominid heroes through the kaleidoscope

Andrew Hill

Missing Links: The Hunt for Earliest Man. By J. Reader. Pp.272. ISBN 0-00-216091-9/0-361-73590-6. (Collins/Little, Brown: 1981.) £9.95, $19.95.
Lucy: The Beginnings of Humankind. By D.C. Johanson and M.A. Edey. Pp.409. ISBN 0-671-25036-1/0-246-11362-6. (Simon and Schuster/Granada: 1981.) $16.95, £9.95.

PICTURE yourself in a camp by a river, in the remote Afar desert of northern Ethiopia. One of the oldest and most spectacular of fossil hominids has just been found; Beatles' music plays loudly over the expedition tape recorder; the fossil — parts of a female skeleton — becomes affectionately known as "Lucy" after *Lucy in the Sky with Diamonds.* The book of the same name, co-authored by Don Johanson, one of the expedition's leaders, is an extravagant romp across the slippery surface of palaeoanthropology. It tells the story of the International Afar Research Expedition and its work, with a bit of palaeoanthropology's historical background, a bit of Johanson's historical background, a bit of gossip. The story is an exciting one, for the expedition has collected large numbers of hominid specimens, including Lucy, which are over three million years old, and probably the most important recently found. Along with similar fossils from Tanzania they are the oldest known hominids from anywhere in the world, and the collection represents a considerable achievement. With a little help from his friends Johanson interprets it in a way he describes as controversial, by suggesting quite plausibly that it all comes from one species, which he has named *Australopithecus afarensis.* This species was dentally and cranially primitive, but was fully bipedal, and he suggests it gave rise not only to the rest of the plio–pleistocene australopithecines, but is also the ancestor of that other prominent hominid lineage — *Homo*, leading to ourselves.

With this material it would be difficult to write a totally uninteresting book, but nevertheless I was disappointed, although initially well disposed to it. Most recent popular books on the subject have come from one factory, and it might have been rewarding to sample another brand. Also the promise of a subjective account was attractive; one that treated the people behind the palaeontology, their personalities, motives and rivalries. After all palaeoanthropology has always been one of those activities where not uncommonly participants gently stab each other in the front, and some consideration of this might have revealed useful information about the dynamics of the science as well as successfully appealing to the general reader. Unfortunately most of the gossip seems gratuitous, or at least simply contrived to reinforce Johanson's ego and status in the discipline, which gradually emerges as one of his chief concerns. Whilst initially amusing, it eventually becomes a little tedious. Similarly the style, with its lapses into chatty dialogue, does not work for me, and at times has a patronizing aspect. Others, however, may find the book a good read, and as it has been well promoted it may do good by interesting the public in a fascinating field. A personalized, subjective account depends largely for its effect on how interesting the author is. Despite its title the real hero of the tale is not Lucy, nor even palaeoanthropology, but Don Johanson himself. If you are interested in Don Johanson, this is the book for you.

If you are interested in palaeoanthropology the book for you is *Missing Links* by John Reader. Unlike Johanson, Reader has no pretensions to being one of the world's leading palaeoanthropologists. He is not a professional scientist, but a very professional photojournalist with a varied career in Africa and elsewhere. His previous book concerned African wildlife; now we have Reader's digest of fossil man.

But this is a serious book and deserves serious treatment. It is the best popular account of palaeoanthropology I have ever read, and simultaneously makes an original and stimulating contribution to the science. Perhaps its most striking feature is Reader's outstanding photographs of fossils, discoverers and memorabilia, but his text matches these. It provides a palaeoanthropological history in terms of its major discoverers and discoveries from 1857 to 1980. This information, which alone is intrinsically interesting, is framed by a theme now common enough in the exegesis of scientific method but which is seldom publicized. Reader indicates that hypotheses are prior to data, and that they are heavily conditioned by social factors, preconceived notions and prejudice. He suggests the acceptability of ideas is closely connected with ". . . the personality and persuasive ability of their proposer. . .", that theories ". . . reveal as much of the

Journals supplement

On October 1 *Nature* will publish a review supplement devoted to new journals. The supplement will cover science journals which have as their initial publication date the period between January 1978 and May 1980, which appear at least three times a year and which use English as the primary language. Publishers and societies are invited to submit four sample issues, including the first and most recent, of journals which fit the above criteria for consideration for review. Journals should be sent to Reviews Editor, *Nature,* 4 Little Essex St, London WC2R 3LF, England.

scientist as of the science . . .". Palaeoanthropology is an eminently suitable field in which to pursue this idea, a field guaranteed to give Feyerabend a good time. Many of its foremost practitioners include people of strong personality and persuasive ability, often well equipped with preconceived notions. Also, like a number of sciences with an historical element, the constraints upon conjecture are relatively slight, for the material evidence for any particular theory is rare, and even more rarely unequivocal. "Controversial" simply means that other people do not believe what you say, and this condition is not new in palaeoanthropology. Each scientist's beliefs can often be justified scientifically, depending for example upon his opinion about time, extent of variation, sexual dimorphism, or the weight given to various anatomical characters. Reader traces these kinds of influences over the past hundred years or so with considerable insight, consideration and knowledge. He does not ridicule people or their beliefs and attitudes, merely suggests that this is the way the science has advanced; and advanced it has despite these strong personal elements in its methodology.

It would be a great pity if *Missing Links* fell into the shadow of other books by perhaps better known or better publicized authors, for it deserves a wide readership. Not only is it an entertaining and informative book for the general public, but it should not be ignored as an anthropology textbook, providing a thoroughly documented history from a viewpoint that more palaeoanthropologists would do well to share. They might remember the girl with kaleidoscope eyes — just shake the evidence and it falls into another pattern, often equally attractive as the one before.

Andrew Hill is a Research Fellow at the National Museums of Kenya and at Yale University.

Understanding the facts of IR spectroscopy

Richard C. Lord

Advances in Infrared Group Frequencies. The Infrared Spectra of Complex Molecules, Vol.2, 2nd Edn. By L.J. Bellamy. Pp.299. ISBN 0-412-221350-3. (Chapman & Hall: 1981.) £15, $35.

IN the past ten years or so there has been a relative decline in the use of infrared spectroscopy for the solution of problems in structural organic chemistry, primarily because of the widening scope of nuclear magnetic resonance and other techniques. For lack of practitioners among recent graduates, skilful interpretation of the infrared spectrum of a complex molecule is becoming something of a lost art. In the preface to this book, the author puts the reason for this in a nutshell: "One cannot interpret an infrared spectrum without a good knowledge of the experimental facts, but too rigid an interpretation without any understanding of how these facts originate can lead to gross errors". Happily, he has provided us with an admirable summary of the experimental facts in successive editions of *The Infrared Spectra of Complex Molecules*. Now, with the second edition of Vol.2 of this work, he gives an up-to-date and comprehensive basis for an "understanding of how these facts originate".

The book is quite compact. The eight chapters cover less than 300 pages and treat successively the alkanes, carbon–carbon double bonds, triple bonds and cumulative double bonds, X–H stretching frequencies, carbonyl frequencies, other double-bond vibrations, the stretching frequencies of XO_2 groups and the effects of hydrogen bonding on the infrared spectrum. Each chapter provides a thorough discussion of mechanical, geometrical and electronic factors that affect the group frequencies under consideration. There are up-to-date references (to 1979) at the end of each chapter, totalling about 1,500 in all.

To me, the chapters on olefinic double-bond frequencies, carbonyl vibrations and hydrogen bonding seem particularly useful. That on carbonyl frequencies is rightly the most extensive; a band due to a carbonyl group is often the most intense in the spectrum of an organic molecule and its precise frequency is a measure of the net effect of a multitude of factors. These include the masses of adjacent atoms, their geometry, interaction with neighbouring vibrations or with overtones (Fermi resonance), conjugation with double bonds and other electronic effects, and intermolecular forces. After detailed discussion of these factors, they are summarized by a table of some 500 carbonyl frequencies in suitably illustrative organic molecules.

This volume gives a thorough account of the various origins of group frequencies in organic molecules. It is a valuable work of reference for those who wish to improve their talents in the interpretation of the infrared spectra of such molecules. Taken with the companion Vol.1, it gives a comprehensive survey of the state of our understanding of the infrared spectra of complex molecules.

Richard C. Lord is the Director Emeritus of the MIT Spectroscopy Laboratory, where he founded in 1950 the first postgraduate training course in the technique and applications of infrared spectroscopy.

Plastid development

F. R. Whatley

Chloroplasts. Results and Problems in Cell Differentiation, 10. Edited by J. Reinert. Pp.240. ISBN 3-540-10082-2. (Springer-Verlag: 1980.) DM 78, $46.10.

THIS book can be read in sequence from cover to cover and represents a beautifully balanced treatment of plastid development and differentiation. All the articles are pitched at the same level and it is easy to cross-refer from one to the other.

In the opening contribution, Schnepf lays down the framework for the rest of the book in describing the development and interconversions of plastids; using evidence from electron micrographs, he provides a classification of all known types. Butterfass next discusses evidence for the continuity of plastids, and also the coordination of the plastid population, both in terms of the origin of the organelles by division and the evolution of patterns of plastid maintenance.

At a more biochemical level, R. G. Herrmann and Possingham give an account of plastid DNA, its amount and arrangement in the circular chromosome; Wollgiehn and Parthier put a similar emphasis on biochemistry in their discussion of RNA in plastids and aspects of protein synthesis. The important chlorophyll–protein complexes associated with light harvesting are then treated in some detail by F. H. Herrmann *et al.* in the chapter on thylakoid biosynthesis, which also includes a discussion of the sites of synthesis of particular components under the control of nuclear and plastid DNA.

Fraction I protein (ribulose*bis*phosphate carboxylase) is discussed by Bottomley, who considers the composition and origin of the component sub-units, and Sundqvist *et al.* rehearse the influence of the main external factors on chloroplast differentiation, especially the effects of light, hormones, minerals and water stress. Finally, the experimental evidence that chloroplasts can be cultured outside the cell is drawn together by Leech, who points out the possibilities of experiments which can be done with surviving plastids.

Regrettably, many of the articles end with a note, added in proof, that the literature for this review was completed in 1976. Since the book was published some four years later the editors must have experienced serious delays in receiving some of the manuscripts. In spite of this, all concerned are to be congratulated on the impressively balanced treatment, which covers structural, physiological and biochemical considerations. As well as being of great help to research workers, advanced undergraduates should also benefit from the book.

F. R. Whatley is Sherardian Professor of Botany at the University of Oxford.

nature

25 June 1981

Congress and cancer research

Has the United States Congress, having declared "war on cancer" ten years ago, now decided to declare war on the National Cancer Institute? This speculation has run through the biomedical research community since the extraordinary occasion three weeks ago when the director of the institute, Dr Vincent T. DeVita, was hauled over the coals by the Senate Committee on Labor and Human Resources (see *Nature*, 11 June, p.444), one of the several committees in Congress which have in the past decade cheered on the institute's growth. The committee was the platform from which, in the early 1970s, Senator Edward M. Kennedy sought to outbid the then Administration in generosity towards cancer research. Now, Senator Kennedy finds himself in the unfamiliar role of senior minority member of the committee, subject to the tutelage of Senator Orrin G. Hatch, apparently bent on winning his spurs in the Senate by being a conspicuous and good Republican — eliminating "waste", diminishing the "role of government" and so on. But Dr DeVita was not merely the victim of a change of political style.

So much seems implicitly to have been acknowledged by Dr DeVita's colleagues at the National Cancer Institute, who put out a statement last week praising his qualities both as a scientist and an administrator. Everybody, at the institute and elsewhere, is agreed that Dr DeVita's appointment as director on 1 January brought a breath of fresh air to an institution in danger of drifting complacently with the tide of congressional largesse. He commands the professional respect of his colleagues, but he is also energetic and tough. It would be disastrous for one of the most important of all federally supported scientific institutions if the Senate committee, or some other zealous custodian of the public purse, were at this stage needlessly to interfere with Dr DeVita's chance of making a success of the National Cancer Institute. The particular issues with which he was confronted three weeks ago — the discovery of irregularities in the conduct of research contracts let by the institute — are acknowledged to have their origins in relatively ancient history. To the extent that the National Cancer Institute is at all responsible, the blame should rest with Dr DeVita's predecessors. His own fault — which some might think a virtue — at the hearing three weeks ago was that of political innocence. He seemed not fully to have appreciated that the first duty of any public servant before a congressional committee is to affirm that his every waking hour is spent on saving every possible nickel of taxpayers' money. Only when that has been done is it permissible for him to think of prosecuting the task for which his budget has been set aside, "curing" cancer or whatever it may be.

So far, there is no knowing what the new Senate Committee on Labor and Human Resources has at the back of its collective mind. The complaints against the institute brought up by Senator Hatch were hardly new. Some of them had been rehearsed before Representative Albert Gore's subcommittee in the House of Representatives some weeks earlier. But, in any case, why should the National Cancer Institute itself be blamed if one of the several investigators working on a collaborative clinical research contract has been accused of keeping inadequate or even falsified medical records? (The investigation of these allegations by the National Institutes of Health will not be complete until after the summer.) Dr DeVita was embarrassed three weeks ago because he had not blackballed a further grant to the same researcher, on the principle of "innocent until proven guilty". But who, in Congress of all places, would expect a good administrator to behave differently? So why did Senator Hatch hold his inquiry at all, and then assiduously leak the details of his complaints in advance to the daily press? Can there be so much smoke and no fire?

Time will tell, probably quite quickly. The unexpected resignation of Dr Donald Fredrickson as director of the National Institutes of Health (of which the National Cancer Institute is one) will be a chance for finding out (see page 603). Although it will be necessary to read the tealeaves before knowing what to make of the White House's nomination of a successor to Fredrickson, Senator Hatch's committee will no doubt seize the opportunity for saying where it stands on continued public support for biomedical research and for cancer research in particular. The greatest danger, in the period immediately ahead, is that the naivety that inspired the war on cancer ten years ago will be found still to persist. Already, some members of the Hatch committee have been asking aloud why, after a decade's generous spending, cancer has not yet been "cured". More generally, the persistence of rheumatoid arthritis, cardiovascular disease or even death itself has lent weight to the popular suspicion that good health is not one of the products of research. Congress, more than usually eager to do the work of the Office of Management and Budget, could easily become jaundiced.

Of necessity, neither obduracy nor round robins in support of Dr DeVita will meet the need. More to the point would be a determined attempt to clear up the misunderstanding between Congress and the research community on the function of cancer research. The notion that there is a war to be won is foolish — and springs largely from the politicians. While the National Cancer Institute has direct responsibility for several important programmes directly related to the improvement of various cancer treatments, the most promising part of its work remains its long-term research, only some of which is ever likely to help with the treatment of specific types of cancer. Congress must somehow be helped to appreciate this simple truth. The task would be easier if it were also more widely appreciated in Congress that the basic research programmes of several of the constituent institutes of the National Institutes of Health are essentially part of a cooperative effort, aimed at the understanding of cell biology and physiology that will, on past experience, make possible as yet unknown treatments for possibly unrecognized illnesses. Administratively, that argues for closer relationships between the several institutes and for flexibility in the disposal of their separate budgets. In the past few years, unfortunately, Congress has been pushing in the opposite direction, seeking more direct control over individual institutes. One consequence is that people like Dr DeVita, appointed as scientists, find themselves dealt with in Congress as if they were nothing but accounting officers. The time has come to reverse this tendency.

No Irish science politics

The assumption that science and politics do not mix is only occasionally falsified. The dangers in too rigorous an attachment to the principle is that it may make a nonsense of the practice. This is what the Science Research Council (now even more augustly known as the Science and Engineering Research Council) has just done. For the council has turned down an innocuous proposal that there should in future be an exchange of graduate students between the United Kingdom and the Irish Republic on the modest scale of a dozen or so a year. By all accounts, the council wishes this modification of pure chauvinism to be subsumed in an

0028-0836/81/260601-02$01.00

even more imaginative scheme — a grand European scheme for the freedom of movement of graduate students within Europe as a whole. The snag, of course, is that the grander scheme will take ages to arrange. Once more, the best is the enemy of the good.

It will not have escaped the attention of the world at large, scientists included, that the United Kingdom government has a serious problem in Northern Ireland, otherwise called Ulster. Some months ago, it was agreed that there should be informal talks on the development of closer relations between the Republic and the United Kingdom between the prime ministers of the two independent states. The talks were known as the "Thatcher-Haughey" talks until the prime minister of Ireland allowed his politicial future to be put in doubt at last week's general election. By all accounts, the agenda for these talks has been severely practical — how to organize the more economic exchange of electricity and so on. The long-term objective has been to build common institutions between the two principal governments of such strength that the people who live in Ulster will appreciate, whatever their immediate inclinations, that the status of Ulster is not the only problem in the world. Somewhere along the line, it now seems plain, somebody had the wit to recognize that a fuller integration of higher education between the United Kingdom and the Republic would serve as a demonstration to Ulstermen of all persuasions (but for practical purposes there are only two) that the interests of young people and of scholarship can transcend sheer sectarianism. The proposal was that up to a dozen places for graduate students in one country might be made available (subject to a proper scrutiny of academic standards and so on) to would-be graduate students from the other. It has been turned down by the Science and Engineering Research Council.

On the face of things, the explanation is reasonable enough. Why make an exception of the Republic of Ireland when the United Kingdom is part of a much larger community? So instead of an understanding between the two native English-speaking governments in Europe on graduate education, the Science and Engineering Research Council is to embark on an exploration of the possibilities of exchanging graduate students more widely, perhaps within the European Community or more widely still. With luck, there may be a formal proposal for everybody to discuss five years from now, by which time the climate of Anglo-Irish relations will have changed, perhaps even for the worse.

This is why the Science and Engineering Council of the United Kingdom should think again. It could reconsider its refusal of a deal on Anglo-Irish graduate students under the item "Matters arising from the minutes" at its next meeting. It should then feel free to acknowledge that it was needlessly pedantic at its earlier meeting, and that there are some circumstances in which the improvement of the social climate that may be brought about by an apparently administrative chance are unexpectedly great. Briefly, the benefits of some kind of understanding between the United Kingdom and the Republic of Ireland on the education of graduate students could suggest to people on both sides of the border between them that there is an even stronger case for running the two university systems in concert, and that still other forms of collaboration might yield still greater benefits. The problem of Ulster is, after all, a problem that the people most concerned have decided is too important to be talked about, let alone negotiated. To those who live in the United Kingdom (including Ulster) and in the Republic, the cruel irony is daily denial of well-established links. The Science and Engineering Council has fallen into the trap of supposing that links such as these are irrelevant to its grander purpose. In that, even in its own self-interest, it is mistaken.

Luddites of two kinds

Luddites are not especially welcome when everybody's energy is concentrated on making national and other (regional? global?) economies more productive. Yet Luddites have a valuable place even if only as the grains of sand that enable oysters to grow pearls. Two interesting but disparate clutches of them have made their appearance recently in the United Kingdom — earlier this week, an organization called Cambridge Econometrics, which is an offshoot of the Department of Applied Economics (at the university of the same name) and which has a well-established reputation for the advocacy of left-wing protectionism, published its latest gloomy economic forecast for the United Kingdom, and the Council for Science and Society, a high-level do-gooder organization last month published its latest document, *New Technology: Society, Employment and Skill.* Both arguments bear on the question whether foreseeable technological developments will make most of us happier than we are. Although different in texture and substance, their conclusions are similar — things will get worse before they can get better; and in the lifetimes of those now alive, the sense of things getting worse will predominate.

Against the odds suggested by its past performance, the Council for Science and Society has the more interesting case. Its document is the product of a committee under Professor Howard Rosenbrock of the University of Manchester Institute of Science and Technology which has obviously had the wit to leave the drafting of its report to one person. The result is a sensitive account of the nineteenth century Industrial Revolution followed by an imaginative guess about the next few decades. The most arresting part of this seemingly rambling train of thought is the brief discussion of the meaning of the word productivity and its suggestion that the concept is not by definition virtuous. If increased productivity means that the same number of people produce more goods collectively, there are more for them to share between them. If, however, increased productivity helps business enterprises to be more competitive, the result may be that they choose to cut each others' throats, not in the process increasing the wellbeing of their workers. Nation-states blessed with increasing productivity have a further degree of freedom — they may choose to spend the extra benefits on research and development or perhaps even on quite different things — armaments for example — but are constrained in the malevolence they are able to cause by social (and political) requirements that, for example, even unemployed workers should be paid nearly a living wage. But in whatever context productivity may be measured, there are potential drawbacks for ordinary people; briefly, they may not share fully in the benefits. This thoughtful document is not a tract but a provoking reminder that more automation — almost certain to come about — does not necessarily mean more happiness.

Cambridge Econometrics, as befits the dismal science that has spawned it, has a narrower perspective. Faced with the long-standing decline in the United Kingdom's performance (and apparently ignorant of the very recent improvement of the terms of trade caused chiefly by the declining value of the pound sterling), the organization comes to the conclusion that unemployment in Britain (expected to exceed 2.5 million this week) will go on increasing for the next ten years at least. The explanation of this odd state of affairs is simple; the British economy will be managed (by economists) as well as may be expected, but the replacement of manual jobs by machines will just keep ahead of the employment of the unemployed in the service industries. The dice, the forecasters say, are loaded against us.

The difference of style between these two gloomy visions of the future is instructive and surprising. The economists, fond as they are of claiming sensitivity to human behaviour, have blundered by pretending to calculate the difference between two large numbers — the number of people displaced from existing jobs plus that of the nett number of those recruited to the labour force — and the number taken up into new industries with an accuracy of less than one per cent. The conclusion that the problem must be made to go away by import restrictions is familiar but irrelevant. The Council for Science and Society is more perceptive, recognizing that it cannot foretell the future with any accuracy, but that the future is likely to be as painful as was the past. That both the future and the present may be pleasurable is hardly mentioned, but that is now the fashion.

0028-0836/81/260602-01$01.00

Fusion researchers face fight for funds

Weapons lobby will oppose Reagan's cuts

Washington

Research into inertial confinement fusion (ICF) — the use of lasers or particle beams to ignite pellets of tritium and deuterium — is fighting for its political life. Supporters in Congress are trying to head off a direct frontal attack by the Reagan Administration, which wants to reduce the ICF operating budget by 25 per cent next year. And already it seems virtually certain that funds will dry up for what some see as the most technically-promising way of obtaining usable energy from such fusion reactions, the use of heavy ion beam drives.

ICF research has been a constant battleground for several years. At present the research is funded as a defence-related programme within the Department of Energy, since the explosion of the pellets can be used to study the simulated effects of nuclear weapons. The Armed Services Commitee in the House of Representatives has been a particularly enthusiastic supporter, championing the construction of facilities at the Department of Energy's two weapons research laboratories and the Sandia Laboratories in Albuquerque, New Mexico, at a total cost of more that $1,000 million since 1977.

However, scepticism within the Office of Management and Budget over whether this level of expenditure is justified has recently been fanned by several factors. One is the continuing difficulty being experienced by the laboratories in approaching the point of thermonuclear ignition, and another is the extreme tight-fistedness of the new Administration, seeking in particular to find ways of reducing federal support for applied energy research.

Things were already turning sour before Mr Reagan came into office. When Mr Carter presented his budget for 1982 to Congress in January, he proposed a total of $219 million to support ICF, the same level of funding as last year with no increase to allow for inflation, and considerably less than the Department of Energy had requested. The Reagan Administration reduced this still further to $180 million. In particular, it suggested that operating funds for the ICF programme be reduced from Mr Carter's proposed $140 million to $106 million, a figure which could result in the lay-off of 300 to 400 scientific and technical staff according to the director of the Office of Inertial Fusion, Dr Richard L. Schriever.

These budget reductions are being sharply contested by the House Armed Services Committee, which is responsible for overseeing the ICF programme. The committee voted to restore all of the money which the Reagan Administration wants to cut, increasing the department's request for operating funds from $106 to $156 million. This represents the single largest increase recommended in the weapons research budget, and was accepted by the full House of Representatives when it voted on the authorizing bill last week. It now passes to the Senate, and then on to the appropriations committees, where it could get a rougher ride.

The Office of Management and Budget is unlikely to give up the fight. There are rumours of demands for even bigger cuts in the ICF budget for 1983. In particular, officials are keeping up their efforts to quash plans supported by the Armed Services Committee to upgrade the NOVA glass laser facility under construction at the University of California's Lawrence Livermore Laboratory.

Scientists at Livermore strongly refute charges that NOVA represents unnecessarily extravagant expenditure. They point to recent research results which have raised hopes that thermonuclear ignition can be reached on schedule by 1986–87. In particular last year's discovery that much higher efficiencies in beam-target coupling can be achieved using shorter wavelength light. This is already being used as the basis for a planned experiment, using two of

Fredrickson quits without warning

Washington

To the consternation of the US biomedical research community, Dr Donald Fredrickson announced unexpectedly last Friday that he will be leaving his position as director of the National Institutes of Health (NIH) on 1 July.

Dr Fredrickson cites "personal" reasons for leaving the world's largest biomedical research organization, referring in particular to the weight of administrative responsibilities and a desire to return to his research in lipoproteins and lipid transport. He has denied any differences of opinion with the new Administration over the funding of biomedical research, although the speed of his departure has inevitably provoked rumours of problems which have yet to surface publicly.

One point of particular frustration to Dr Fredrickson is said to have been the difficulty of making appointments to top positions at NIH at a time when any such decision requires a special exemption from the current freeze in federal hiring. Several senior administrators have recently left NIH, including Dr Robert L. Goldberger, deputy director for science, who has been appointed provost and vice-president for health sciences at Columbia University in New York.

A respected scientist and a popular administrator, Dr Fredrickson first joined NIH 28 years ago, and was appointed their director by President Gerald Ford in 1975. He had previously been director of the National Heart and Lung Institute, and in 1974 became president of the Institute of Medicine at the National Academy of Sciences.

Dr Fredrickson is said to have a good working relationship with the new Secretary of Health and Human Services, Mr Richard Schweiker. Mr Schweiker had previously been a keen supporter of NIH as a member of the Senate's Human Resources Committee, and is said to have recently agreed to reconfirm Dr Fredrickson's appointment — just as Mr Joseph Califano had done when Mr Carter became president in 1977.

If Dr Fredrickson is having problems with other members of the new Administration, then they are being well concealed. The conviction that there must be some unexplained reasons for his rapid departure seems partly an expression of disappointment that a popular director is leaving NIH in the hands of an unknown successor just when possible budget restrictions could mean a new period of uncertainty. During Dr Fredrickson's tenure at NIH the research budget has increased from $1,880 to about $3,500 million, although the rate of expansion has lately slowed considerably.

No whisper where next

Last week, in delivering an emotional address to NIH staff at the end of which he announced that he had submitted his resignation to President Reagan, Dr Fredrickson said that NIH were "healthy, strong and in the prime of life", while warning that the essence of the institutes' greatness was "fragile and could be destroyed by careless trustees". He has made no announcement about future plans, apart from telling senior staff that he will spend some months with the National Academy of Sciences. **David Dickson**

0028-0836/81/260603-02$01.00

NOVA's beams and equipment from the Argus and Siva lasers — an experiment known as NOVETTE — operating with green light. This experiment will make it possible to test parts of NOVA before the full facility becomes operational, a cost-reducing exercise which, it is hoped, will please the sceptics at the Office of Management and Budget.

Contingency plans are also being made at the Los Alamos Scientific Laboratory, where the 35–40 kilojoule carbon dioxide laser ANTARES is due to be completed in 1983, and at Sandia Laboratories, where construction of the first phase of a new particle beam fusion accelerator has just been completed. Each of these programmes would be severely affected if the full cuts proposed by the Reagan Adminstration go through.

Given the political strength of the Armed Services Committee, and the general enthusiasm for military-sponsored research, it seems unlikely that the full cuts will, in fact, materialize. There is much enthusiasm for the military use of lasers among conservative Republicans, and the White House's National Security Council is studying a proposal for substantial increases in funding for a space-based, high-energy laser system, and has given support to a decision by the Senate to add $50 million for research in this area.

A more likely casualty of the budget stringencies, however, is heavy ion beam research. From a purely technical point of view, many experts believe that the use of heavy ions to explode (or more accurately, implode) the pellets — as opposed to the light ions being studied at Sandia Labs — present engineering difficulties in designing a reactor. But heavy ions continue to suffer from two major drawbacks.

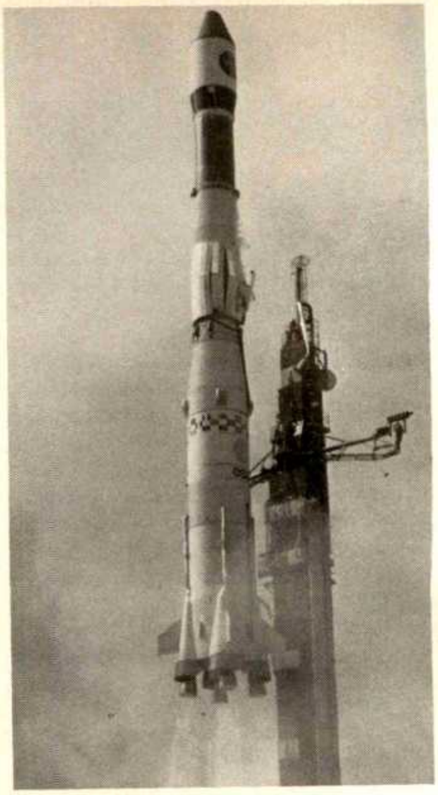
First, it would cost many hundreds of millions of dollars to build an experimental machine, expenditure that stands little chance of gaining federal support in the prevailing financial climate. And second, unlike either lasers or light ions, such technology has no potential military uses.

Support from the Armed Services Committee has ranged from lukewarm to frigid. And given its own priority of glass lasers, it has given heavy ions a firm thumbs down, not only supporting the Administration's decision to withdraw funding for this area of research (which received $5 million in the current year), but adding explicit language in the authorizing bill forbidding the department to provide support from other funds.

Supporters of heavy ions had hoped that their cause would be strengthened by a review of ICF previously announced by the Department of Energy's Research Advisory Board (*Nature* **287**, 573; 1980). But with the change in administration this review has been put on ice; and with both Congress and the Department of Energy now stressing that this principal interest in ICF lies in its military applications, the prospects for heavy ion drivers look increasingly bleak. **David Dickson**

European space success

Ariane takes off



Kourou, French Guiana

Champagne corks were popping here last Friday after the near perfect third test launch of Ariane, Europe's hope for an independent satellite launcher. During the final countdown, tension was high amongst the largely French team of engineers who had worked to correct the fault in the first stage Viking engines that had led to the failure of the second test launch a year ago. They, at least, are now reassured that the earlier problems have been ironed out.

Lift off was at 12.33 GMT, 19 June, on the third countdown, about 1 hour 10 minutes later than scheduled. The first two attempts were abandoned after a marginal overload in the power supply to the third stage, and when two local radar stations failed to synchronize. In the event, the voltage tolerance was increased and only one radar station tracked the rocket during the first few minutes of flight. Preliminary analysis of data indicated that the trajectory and timing were on target and that the three payload satellites, Meteosat II, Apple — the Indian experimental communications satellite — and a technological capsule, were injected into a suitable geostationary transfer orbit.

Over the next few weeks, detailed analysis of flight data will reveal precisely how the first stage engines performed. The worry had been that high frequency combustion instabilities at 2,300 Hz and 2,700 Hz, which had resulted in the failure of the second launch, might recur. Although the mechanics of the instabilities, not uncommon in liquid fuel rocket engines, is not understood, the problem seems to have been solved largely by redesigning and increasing the size of the orifices in the hydrazine and nitrogen tetroxide fuel injectors.

With two successful test launches behind it, Ariane is now qualified for commercial operation, according to the European Space Agency (ESA). After the fourth and final development launch, scheduled to put ESA's maritime satellite, Marecs, into orbit on 10 November, commercial operation will be handed over to Arianespace, a consortium of European companies. The fourth development flight, which will launch within a particularly narrow window at night, will be a more stringent test of Ariane's capabilities than last week's exercise. Subsequent commercial launches will confirm whether the launcher achieves its design reliability of 95 per cent.

Confidence, however, is now such that work began earlier this week on building a second launch pad at Kourou, finance for which is expected to be approved by the ESA member states later this month. ESA is hoping that last week's success will prompt enough orders to fill the ten launch per year capacity of the two pads. So far seven firm orders and fourteen reservations have been made, mainly for placing communications satellites into geostationary transfer orbit. The hope is that two or three more firm orders will be placed by the autumn and that several customers, including the Arab countries, which will want to launch two Arabsat telecommunications satellites in 1983 and 1984, and the Australians, will now be prompted to opt for Ariane rather than the United States space shuttle, which, despite its successful inaugural flight, still remains to be qualified.

Meanwhile, ESA's role in Ariane's development is not yet over. Work is under way on Ariane 2 and 3 with payload capacities of 2,000 kg and 2,400 kg respectively, compared with the present capacity of 1,700 kg. Hopes are now high within the Centre National d'Etudes Spatiales (CNES), the French space agency, that the ESA member states will be sympathetic to Ariane 4 with 3,500 kg or payload capacity. CNES is also planning to propose a partly reusable Ariane 5 later this year.

The success of the French ideas, however, will depend on how the other member states decide that ESA's budget should be spent over the next five years. Negotiations so far have ended in deadlock with the three largest contributors, France, Germany and Britain, wanting to go their own ways. But it is not surprising that after contributing more than 60 per cent of the money and effort spent on Ariane development, France has claimed the lion's share of last week's success and is particularly anxious that the programme should not stop there. **Judy Redfearn**

0028-0836/81/260604-01$01.00

Soviet space research

Salyut era ends

Despite news of last Friday's docking of the unmanned Kosmos-1267 probe with the now-empty Salyut-6 space station, the Soviet media are now celebrating the end of the Salyut-6/Soyuz programme and Mr Brezhnev has outlined the next stage of the Soviet space effort, permanent orbital stations served by relays of crews.

Mr Brezhnev was speaking at an Order of Lenin investiture ceremony for cosmonauts Kovalenok and Savinykh, just returned from Salyut-6. The projected permanent space station, which, following Tsiolkovskii, Soviet space planners see as a necessary precursor to manned flights to the Moon and planets, will, said Mr Brezhnev, involve "more complicated tasks" — an indication that that will probably be a considerable delay before such a station is fully operational. However, the docking of the two unmanned vehicles Kosmos-1267 and Salyut-6 is clearly a step towards working out efficient ways of constructing large orbiting complexes.

The achievements with Salyut-6 have been considerable: 28 cosmonauts visited the station including participants from the Comecon partners, there were 34 manned and unmanned dockings, and about 900 separate biological and medical experiments. Some cosmonauts have remained in zero gravity conditions for as long as six months, and the responses of man to prolonged weightlessness will have been well studied.

Before pressing on with new orbital stations though, there is some unfinished business for Salyut planners. Both France and India have been promised berths for their own cosmonauts on Salyut-type craft, and the Soviet news agency Tass has already reported details of the proposed Franco-Soviet medical programme which is to include a new French *krvotock* (Russian for "blood-flow") device which only needs to be pressed against a cosmonaut's neck to allow the associated computer to record all the necessary details about cerebral blood flow. **Vera Rich**

Telescope gets director

Professor Riccardo Giacconi, associate director of the Harvard-Smithsonian Center for Astrophysics, has been appointed the first director of the Space Telescope Institute to be established at Johns Hopkins University in Baltimore, Maryland. The appointment, which takes effect on 1 September, was announced last week by the Association of Universities for Research in Astronomy (AURA), which earlier this year won from the National Aeronautics and Space Administration the contract to manage the new institute.

Professor Giacconi has been an X-ray astronomer since he joined American Science and Engineering in 1958. Over the past few years, at the Harvard-Smithsonian Center for Astrophysics, he has been the scientific and administrative leader for the highly successful Einstein X-ray astronomy satellite. The new institute should be built and staffed by 1982; the launch of the space telescope is scheduled for 1985.

European participation in the project will ensure at least 15 per cent of observing time for European astronomers who may also be represented at the same sort of level among the institute's staff. The European counterpart of the Johns Hopkins institute, the proposed European Co-ordination Facility, is now likely to be sited at Garching, near Munich. That at least is the recommendation put forward by a European Space Agency panel set up to investigate the proposed sites. Alternatives under consideration are the Institute of Space Astrophysics at Frascati, the Royal Observatory at Edinburgh and a joint proposal from the Observatoire de Paris and the Institut d'Astrophysique in Paris. The final decision is expected this week.

Philip Campbell

Endangered species

Belgium joins up

Brussels

The Belgian parliament has now voted to bring Belgium into line with nearly all the other members of the European Community by ratifying the Washington Convention concerning trade in endangered species.

Effectively, this means that at every entry point into the Community, customs officials are now required to ensure that any imports of flora or fauna, or their products, are not listed in the Washington Convention.

The only two EEC countries still to ratify the convention are Luxembourg and the Netherlands. The Netherlands has its own strict legislation and like Luxembourg was only waiting for the other members of the Benelux Union to ratify. The European Commission is trying to have a decision adopted that will allow the Community to speak with one voice.

Legislation is all very well, but there are fears that even now that the last gap has been plugged the demand in Europe for fur skins, alligator handbags, leopard tooth jewellery and the like will ensure a continuation of the trade, for Europe is the biggest market for these products.

Belgium is well known as the European centre of the trade in endangered species. And although imports of animal and reptile skins in 1980 showed a decline for the first time in eight years, they were still worth 234 million Belgian francs (£4 million) according to official figures. Experts at the Belgian section of the World Wildlife Fund think that this is a gross underestimate and that the market value of these skins is likely to be more than ten times as high.

Belgium became the centre for the international trade because, over the past eight years, it has provided the easiest means by which goods could be brought into and distributed around Europe free of further customs controls. The new ruling will, however, require that documentation be presented proving that the animal or animal product in question is required for scientific purposes or for a zoo.

This in itself would be difficult to implement, but the real problem is that most customs officials will not have the knowledge or the motivation to tell, say, an endangered species of canary from an unprotected species and 30,000 vegetable species which it divides into three groups — species in danger of immediate extinction, those threatened with extinction and species in danger in a certain country. Specimens in the first list must be accompanied by both an import and an export permit while those in the second list merely require an export permit.

The first list has become considerably longer since the convention was signed in 1973, and at the Delhi meeting of the 68 member states this year a new proposal was made. Instead of listing those species in danger, the reverse should be done. "Reverse listing" would mean that animal and plant traders would have to prove that their goods are a species listed as safe.

Jasper Becker

UK nuclear power

Marshalling forces

Britain's floundering plans to build a series of pressurized water nuclear reactors are to get a thorough shake-up — with the appointment of Dr Walter Marshall, currently chairman of the United Kingdom Atomic Energy Authority, as coordinator of the programme. The appointment will be formalized this week.

Key technical decisions are falling behind schedule, and although there is no present pressure on electricity supplies the government is said to be worried that the forthcoming pressurized water reactor (PWR) inquiry will fall too close to the next general election, complicating the central economic and social issues with the vexed politics of nuclear power.

Dr Marshall's role will be to ensure that the necessary decisions are taken when they need to be taken, drawing together the confusing interests and bureaucracy of the National Nuclear Corporation (NNC, which will construct the plants), the Nuclear Installations Inspectorate (NII, concerned with safety), and the customers, the Central Electricity Generating Board (CEGB).

The situation has been greatly confused

0028-0836/81/260605-02$01.00

lately by a number of factors: demands by the Installations Inspectorate that British PWRs be better contained, and more free of radiation on the reactor floor, than their American counterparts; worries in the CEGB that the National Nuclear Corporation might be undercapitalized at its present £10 million; the resignation of the director of the corporation (the post remains vacant); and the publication of parliamentary reports on the government's nuclear programme and the CEGB, both of which were critical of the way in which the need for the programme had been assessed.

Marshall has apparently begun by asking the several interested parties to set down on paper a list of the decisions they consider need still to be made, noting those on which everybody has already agreed. He plans then to draw up a timetable for decision on the outstanding questions, confining his own role to making sure that the timetable will be met. Although the Atomic Energy Authority has a substantial programme of research on the safety of the PWR, Marshall plans to be a judicial chairman, not allowing his own organization as such to have the last word.

Dr Marshall, it is hoped, will fill the managerial vacuum without direct responsibility for the technical decisions themselves, which will remain in the hands of NNC, NNI and CEGB. However it seems inevitable that Marshall, with his forceful and argumentative personality, will stamp his mark on the project.

Robert Walgate

Basel Immunology Institute

Dancing on

Basel

The Basel Institute of Immunology, internationally distinguished for its contributions to fundamental research, last week celebrated its tenth birthday with a series of general scientific talks intended to be intelligible to a wide audience, and a ballet intelligible only to immunologists.

The celebration came at a critical point in the life of the institute, whose directorship passed last summer from its founder, Niels Kaj Jerne, to Fritz Melchers, one of its earliest members (see *Nature* **286,** 3; 1980). This has inevitably raised the question of how far the success of the past ten years has been due to the personality of the founding director, expressed both in the organization of the institute and in its intellectual tone. Apart from about ten permanent members, most of whom have been there from its inception, the institute has no tenured personnel: the remaining 50 or so scientists hold two- or three-year contracts and usually leave at the end of them. During their term at the institute, all scientists have strictly equal status.

The other conspicuous feature of the institute under Jerne was the distance kept between the research staff and that of Hoffman La Roche, the Basel-based pharmaceutical company that supports it. Melchers promised when he took over that the distance would decrease, and indeed Roche laboratory personnel do now visit the institute for discussions with its members. So far, however, much the most effective link between the institute and the company has been Theo Staehelin, now working for Roche, but still temperamentally part of the institute.

A more difficult question is posed by the imminent departure of Susumu Tonegawa for the Massachusetts Institute of Technology after ten years in Basel. In 1976 he was the first to announce the sequence of genomic DNA coding for an antibody molecule and has since been consistently among the leaders in the analysis of cloned immunoglobulin DNA. The question of his replacement is problematic in several ways. Research on cloned DNA is now so brutally competitive, and its practitioners so much in demand, that the loss of Tonegawa may also lose the Basel Institute its lead in the molecular biology of antibody genes. But an attempt to maintain such a lead might in any case have dangers for the institute. It would absorb a substantial proportion of the support provided by Hoffman La Roche, and would almost certainly entail the contravention of the institute's principle of equality and autonomy of all its members.

Melchers, however, recognizes that to abandon research on recombinant DNA altogether is out of the question. Although most of Tonegawa's collaborators are leaving either with him or for other academic laboratories, Hiroshi Sakano remains for about a year and will be joined by one of Tonegawa's earlier collaborators. Melchers is also seeking another molecular biologist who will combine expertise in recombinant DNA technology with a feeling for immunology. Immunoglobulin, he points out, is not the only molecule synthesized by lymphocytes — they also produce important intracellular signalling molecules about which almost nothing is known. Research on membrane proteins is another line Melchers would like to see at the institute.

This emphasis partly reflects Dr Melchers' own research interests, which have been in the clarification of the cellular interactions that trigger antibody production by specialized plasma cells, the subject of one of the three scientific sessions the institute anniversary celebration. The other two dealt with immunoglobulin and the major histocompatibility antigens. One topic conspicuous by its absence was the "idiotypic network" proposed by Jerne to explain the regulation of the immune response and which has long been among the research preoccupations of the institute. Jerne, summing up the meeting, commented only mildly on the omission.

The omission was made good in the ballet, performed with astounding verve at the end of the celebration by respected scientific members of the institute and their technical staff, ingeniously costumed as a variety of lymphocytes and associated cells.

Particularly notable were performances from Harald von Boehmer, benignly engulfing as a macrophage; Kirsten Fischer-Lindahl, chilling as the male-specific T killer cell, Ivan Lefkovits clinically efficient as a T helper cell, and, at a more molecular level, Charles Wood as a Samurai recombinase cleaving a *corps* of nucleotides that athletically recombined to code for antibodies too numerous to mention.

Miranda Robertson

German nuclear reprocessing

Decision on site

Attempts are under way to break the stalemate in West German nuclear politics. DWK, the German association for reprocessing spent nuclear fuel, has now announced its choice of a site for a small reprocessing plant on the borders of North Hesse and North-Rhine Westphalia. And the federal ministers for research and the interior expressed the view that there are no technical obstacles to a high-level waste storage facility in the salt domes of Gorleben in Lower Saxony.

Those statements are significant because, if followed by the appropriate action, they could ease the burden on utilities applying for licences to build new nuclear plant to satisfy the government that they can safely dispose of nuclear waste. Since 1976, various methods of waste disposal have been a condition of construction and commissioning licences. Hopes that those conditions could be met through reprocessing and final waste disposal in a geological formations were dampened in 1979 when planning permission for a major facility at Gorleben was refused after a lengthy public inquiry.

Since then, storage tanks at power stations have been upgraded to store for longer and permission has been sought to build intermediate storage tanks — which the inquiry had not ruled out — at Gorleben. The search has also been on for alternative sites for reprocessing plants in states whose governments are known to be sympathetic to nuclear power.

So far, DWK has made investigations in two such states: North Hesse and Rhineland-Phalz. The Rhineland-Phalz negotiations are still at an early stage. The idea is to build two small reprocessing plants each capable of 300–350 tonnes throughput a year. A third plant, to be built later at an as yet unspecified site, would have a 700-tonne capacity bringing the total capacity up to the 1,400-tonne originally planned for Gorleben.

DWK's choice of a specific site in North Hesse means that formal application can now go ahead. But permission will not be granted easily. Although the state government under the leadership of Herr

0028-0836/81/260606-02$01.00

Holger Boerner is sympathetic, opposition within his ruling Social Democratic Party could scupper the scheme. Last weekend, however, Herr Boerner went to his local party for a vote of confidence in his leadership. His mandate as premier of North Hesse was supported, but opposition to the reprocessing plant together with a plan to extend Frankfurt airport was voiced. The likely outcome now is that the reprocessing plant will go-ahead but local opinion will be firmly divided.

Meanwhile, the local inhabitants of Gorleben have reopened discussion on whether a high-level waste storage facility should be built in the area. At a public meeting in May, several objections to the technical feasibility of the facility were aired: the salt domes are not covered by a layer of clay, they are in contact with ground water in some places, the region may not be as geologically stable as originally thought and natural gas is extracted from the eastern section of the domes in East Germany. Nevertheless, the federal government seems convinced that the Gorleben salt domes offer the best site within West Germany and has said that, even if conditions are not ideal, problems can be overcome by tailoring the waste to suit the geological formation. The government is still committed to an operational storage facility by the 1990s and Gorleben still looks like the best option. **Judy Redfearn**

European environment

Win some, lose some

Luxembourg

Environment ministers of the European Community are in the throes of attempting to draw up regulations on the discharge of noxious substances, outlined in the 1976 directive ENV 131. On 11 June, they again came within a hair's breadth of agreement on the seemingly intractable problem of common standards for the discharge of mercury into the aquatic environment. But one outstanding issue remains.

The United Kingdom is continuing to defend its right to apply qualitative objectives for the discharge of mercury from both old and new plants, while Italy and France insist that limiting values should be obligatory for assessing pollution from new plants. The British, it is maintained, are defending a point of principle while their continental partners are protecting their commercial interests. Nevertheless, British officials point out that the Italians, for instance, would find it hard to enforce the proposed maximum values. It is hoped that the disagreement can be resolved within the next three months.

There was further agonizing about the council's decision to postpone the adoption of the so-called Seveso directive on the prevention of major industrial accidents. The French have held up agreement, pleading that the new government has not had time to adopt a fresh policy, but hinting that the Mitterrand regime will be more sympathetic than its predecessor to the outstanding problem of transfrontier notification.

Another subject for debate is the threat to the ozone layer posed by chlorofluorocarbons (CFCs 11 and 12). At the end of this year, the European Community is due to review its commitment to reduce the use of these materials in aerosols to 30 per cent of 1976 levels. The commission's report is unlikely to satisfy either side in the dispute.

Although sales of CFCs 11 and 12 as aerosol propellants fell in 1980 to 28.2 per cent of 1976 levels and are expected to drop to 35 per cent this year, sales for other uses such as refrigeration and plastic foam have risen by 34.7 per cent. The commission's report implies that the situation is well in hand, but favours reducing the production of CFCs for uses other than in aerosols. Negotiations are also to be pursued within the framework of the United Nations Environmental Programme and the Overseas Environmental Council for Development. The council seemed content to assume that, even if the ozone layer is being destroyed at an annual rate of 5–7 per cent current measures are adequate. Even so, the Community appears to be taking a daring step; there is nothing like a consensus among its members that the present restrictions are necessary.

Surprisingly, the United Kingdom is pushing for stringent Community restrictions on lead pollution, on the lines recently adopted in Britain, which would mean a lowering of the lead level in petrol from 0.4 to 0.15 grammes per litre. West Germany and Denmark will be enforcing this limit, and it now seems likely that the Commission's 1976 proposed directive on the permissible level of lead in air will also be taken off the shelf. The council has agreed in principle to endorse an environmental standard of 0.2 milligrammes per cubic metre of air, and has directed that the developer should shoulder the costs of the assessment, that the information should be publicly available and that there should be consultations with neighbouring states. However, 130 reservations were tabled against the directive at the beginning of the council meeting.

With an air of embarrassed pride, the ministers also announced agreement on two measures designed to prevent further disasters similar to that of the *Amoco Cadiz* — drawing up a compendium of the properties of hydrocarbons and a list of oil-fighting equipment in the Community. The more important proposal to keep a register of oil tankers most at risk was resisted by the Greeks and passed on for consideration by the transport ministers. The two proposals agreed are the only measures produced since the European Community announced its crusade against oil tanker disasters following the 1978 *Amoco Cadiz* shipwreck. **Jasper Becker**

Water works for the Cuna villagers

Panama

For the Cuna indians of Panama progress still means hard physical work. But as a result, the inhabitants of the whole Cuna village can now enjoy clean running water.

The Cunas are colourful tribes of interbred indians who live in palm and bamboo huts on the beautiful San Blas islands — a cluster of over 360 small coral islands off Panama's Caribbean coast. The location allows an almost completely autonomous existence.

The island village of Ticantiki, with its 900 inhabitants, is situated on Devil Cays, only 600 yards from the mainland. It is this proximity to the mainland that has allowed fresh water to be piped under the sea to the island.

A new dam has recently been built up in the hills of the mainland as part of a government project involving the labour of the local indians. In indian terms, the dam is three hours' walk through the jungle from the coast. From the dam, fresh water flows through plastic piping down to the island at a rate of 60 gallons per minute.

The building of a water filter has been left entirely to the indians. The filter is situated a half-hour's walk from the coast, about 2 miles. The shell of the filter is made of reinforced concrete and comprises two bins, each about 8 ft square and 6 ft deep. Carefully graded stones are being used to fill the bins and provide the filter surface. Water is sprinkled from overhead through holes in the plastic piping. The filter, though simple in concept, is effective. When it is finished the indians hope to be able to increase the water flow to 80–90 gallons per minute.

But one may ask, where in a tropical forest do you find enough stones for a filter of this size? The answer is, at the mouth of the river. Here, children help in the work by grading stones. The men of the village carry the stones up to the filter in sacks suspended from a pole on their shoulder. This load weighs about 50 kilos, yet these tough, stocky indians stride through the jungle as though it were no more than a sack of cotton.

The congress of the village dictates that any able-bodied man who does not carry up a load should pay $5 a day to the community. For a culture that lives by fishing, tending plantations and, increasingly tourism, this is a large sum of money.

And the next project for this ambitious little village? To pipe their fresh water to every household. **Patricia Dent**

0028-0836/81/260607-01$01.00

CORRESPONDENCE

Huxley on *Nature*

SIR — You quote the president of the Royal Society (*Nature* 4 June, p.373), speaking at the centenary celebrations for the British Museum (Natural History), as criticizing *Nature* for having "printed more than 30 letters . . . . relating to cladistics, which would have been unintelligible to 99 per cent of the readership".

But, sir, would that not apply equally to many of the other topics dealt with, and rightly dealt with, in the pages of your esteemed journal? Elitist it may be, but it is to be hoped that you will continue to give space to technical scientific matters, especially when these are a subject of controversy, commensurate with their interest and importance, even though this may be apparent to no more than a small minority of your readership.

C. B. GOODHART

University Museum of Zoology,
Cambridge, UK

Cabal of referees

SIR — The article in *Nature* referring to the policing of the publication of scientific literature (9 April, p.433) raises a major issue.

We suggest that the main reason for the problems now surfacing and referred to in this article is the existence of a refereeing system which has turned into a cabal and like all secret societies is plagued by corruption.

This problem has already been raised by an assistant editor of the journal *Nuclear Physics*, who felt that the present anonymous system of peer review results in injustices to authors and is inappropriate to the spirit of science (P. Robertson *New Scientist* **71**, 410; 1976.)

Quite clearly, the reputation of this system is now at stake. Could it be that the scientific community at large is not yet mature enough to handle the problems that may arise from a more open system?

M.A. GILLMAN
F.J. LICHTIGFELD

Medical School,
University of the Witwatersrand,
Johannesburg, South Africa

Book for safety

SIR — In an editorial on nuclear politics (*Nature* 4 June, p.368) you suggest that governments should be responsible for safety and responsive to the views of the electors. This is to recognize the democratic imperative, but you might have added the need to inform the electorate about safety matters. With these, as with all difficulties, the strength of a democracy depends on the sense of its citizens.

One of our functions at the National Radiological Protection Board is indeed to provide information on radiation safety, and we do so in several ways. Because of public anxiety about nuclear issues, we discuss them impartially in a new general publication *Living with Radiation* (Her Majesty's Stationery Office, London, 1981; 50 pence) which we hope will promote cool debate and calm judgement.

M.C. O'RIORDAN

National Radiological Protection Board,
Chilton, Didcot, UK

Wheel re-invented

SIR — The recent report by K.S. Jayaraman (*Nature* 28 May, p.277) describing the "gunny-bag air conditioner" developed by the Central Building Research Institute in India, provides an excellent example of the redevelopment of wheels at government expense. The process of passive cooling by roof-surface evaporation is as Jayaraman says suitable not only for India and other developing countries in the tropics but for large areas of American suburbia where for many years the advent of summer has been marked by the temporary placement of lawn sprinklers on the ridge of the roof.

J.P. BENTLEY

University of Oregon Health Sciences Center,
Portland, Oregon, USA

Doomed subspecies

SIR — How would a sociobiologist explain the existence of anti-sociobiologists?

Undoubtedly, most people in the world would claim to be basically generous, caring, and unselfish individuals. Sociobiology, with its fundamental emphasis upon the selfishness of human nature, thus represents a minority view.

But even if the majority view of human nature is untrue, being the result of genes which have been selected for such "deception", might anti-sociobiologists nevertheless receive greater fitness benefits from this majority?

Might one expect to see the demise of sociobiology over evolutionary time due to the decreased fitness of sociobiologists relative to anti-sociobiologists?

ARTHUR MELTZER

Syracuse, New York, USA

Divergent usage

SIR — I would like to note a common and unfortunate misuse of terms. In his *Origin of the Species,* Darwin noted that similarities among organisms could arise from *convergence* because separate lines of descent independently acquired a function, or by *divergence* from a common ancestor. I do not know who coined the words, but Darwin used *analogy* to describe the former and *homology* to describe the latter. That useful distinction has served biologists well for over a hundred years.

The clarity of these concepts is now being widely abandoned as illustrated by your associate editor who presents the conclusion "that the homology is due to divergent rather than convergent evolution"[1]. She thus presents us with a choice between a tautology (homology due to divergence) and a contradiction in terms (homology due to convergence).

One should use *similarity* for the meaning that homology must have in the above quote for reasons far more important than a personal preference for Latin over Greek-derived words or preserving a time-honoured usage. Similarity is an *observation* while homology is an *inference* about its origins. To use homology with both meanings is to confuse the data with the hypothesis. And how much simpler it is to conclude, "The similarity is homologous rather than analogous".

WALTER M. FITCH
RONALD L. NIECE

Department of Physiological Chemistry,
University of Wisconsin-Madison,
Madison, Wisconsin, USA

---

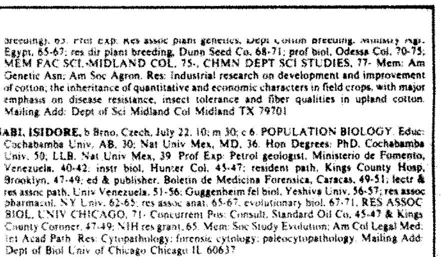
1. Robertson, M. *Nature* **290**, 187 (1981).

Credit due to Nabi

SIR — It has recently been suggested in the columns of *Nature* that I am the mysterious Isidore Nabi (*Nature* 23 April, p.623). I would like to do what I can to clarify the situation. Let me state categorically that any assertion that Isidore Nabi is none other than R. C. Lewontin is incorrect. Let me offer a few corroborative details.

(1) According to his biography in *American Men and Women of Science,* Dr Nabi is 71 years old, received his bachelor's degree from Cochabamba University, and, among other things, has lectured and carried out research at the University of Venezuela for five years. I, on the other hand, am 52, have never even heard of Cochabamba University, and have never been south of Mexico City.

(2) Dr Nabi is the editor of the journal *Evolutionary Theory* on whose editorial board I also appear by name, and I also find him listed as a member of the Evolution Society of which I once had the honour of being president. Why would Professor Van Valen, managing editor of *Evolutionary Theory,* list me on the editorial board if I were also, under a different name, editor of that worthy journal? And what in the world would I do with an extra copy of *American Naturalist* when I hardly know what to do with my own?



Egypt, 65-67; res dir plant breeding, Dunn Seed Co, 68-71; prof biol, Odessa Col, 70-75; MEM FAC SCI-MIDLAND COL, 75-, CHMN DEPT SCI STUDIES, 77- Mem: Am Genetic Asn; Am Soc Agron. Res: Industrial research on development and improvement of cotton; the inheritance of quantitative and economic characters in field crops, with major emphasis on disease resistance, insect tolerance and fiber qualities in upland cotton. Mailing Add: Dept of Sci Midland Col Midland TX 79701

NABI, ISIDORE, b Brno, Czech, July 22, 10; m 30; c 6. POPULATION BIOLOGY. Educ: Cochabamba Univ, AB, 30; Nat Univ Mex, MD, 36. Hon Degrees: PhD, Cochabamba Univ, 50; LLB, Nat Univ Mex, 39. Prof Exp: Petrol geologist, Ministerio de Fomento, Venezuela, 40-42; instr biol, Hunter Col, 45-47; resident path, Kings County Hosp, Brooklyn, 47-49; ed & publisher, Boletin de Medicina Forensica, Caracas, 49-51; lectr & res assoc path, Univ Venezuela, 51-56; Guggenheim fel biol, Yeshiva Univ, 56-57; res assoc pharmacol, NY Univ, 62-65; res assoc anat, 65-67, evolutionary biol, 67-71, RES ASSOC BIOL, UNIV CHICAGO, 71- Concurrent Pos: Consult, Standard Oil Co, 45-47 & Kings County Coroner, 47-49; NIH res grant, 65. Mem: Soc Study Evolution; Am Col Legal Med; Int Acad Path. Res: Cytopathology; forensic cytology; paleocytopathology. Mailing Add: Dept of Biol Univ of Chicago Chicago IL 60637

NABIGHIAN, MISAC N, b Bucharest, Rumania, Dec 5, 31; US citizen; m 66; c 2. GEOPHYSICS. Educ: Mining Inst Bucharest, BS, 54; Columbia Univ, PhD(geophys), 67. Prof Exp: Geophysicist, Geol Comt, Bucharest, 55-57; res scientist, Geophys Inst,

(3) Isidore Nabi is the author of several important works which, I am sorry to say, are not at all of my creation. I refer in particular to his brilliant *"On the Properties of Motion"* which is as yet unpublished but widely circulated and known, and his seminal work *"An Evolutionary Interpretation of the English Sonnet".*

I have recently received a letter from Professor Van Valen saying that *he* has been identified as the Isidore Nabi who wrote the letter to *Nature,* an assertion which he denies. Thus, confusion multiplies. I hope that this letter has thrown some light on the situation.

RICHARD C. LEWONTIN

Museum of Comparative Zoology,
Harvard University, Massachusetts, USA

Cosmic quadrupole?

from Richard A. Muller

"THE truly marvelous thing about science is the great return in theory that one obtains for such a meager investment in fact." — Mark Twain.

In no other field of science is Mark Twain's perception as true as in cosmology. There are so few 'facts' that constrain cosmology that the discovery of a new one is a major event. One of the most recent such discoveries is the anisotropy (variation in intensity with direction) in the cosmic microwave background radiation. It is now well established[1,2] that the intensity of the radiation varies by about 0.2 per cent (6 mK of the 3K signal), with maximum intensity towards the Virgo cluster of galaxies, and minimum in the opposite direction. The intensity of the modulation varies smoothly between these extremes as the cosine of the angle in the sky, suggesting that the anisotropy is due to a Doppler shift from motion of the Milky Way galaxy with respect to the radiation. This cosine modulation has recently been observed[3,4] to have a blackbody spectrum in the frequency range from 19 to over 100 GHz. This observation has eliminated lingering fears that the anisotropy might be a spurious effect due to interference by galactic synchrotron emission.

Two groups are now reporting a new 'fact' — a second harmonic in the cosmic background anisotropy. R. Fabbri, I. Guidi, F. Melchiorri and V. Natale[3] in Italy have observed a second harmonic with amplitude 0.9 (+0.4, −0.2) mK and axis consistent with that of the first harmonic. S. Boughn, E. Cheng and D. Wilkinson[4] at Princeton report a 'quadrupole moment' Q_5 (that is, a second harmonic proportional to $\cos^2\delta \cos 2\alpha$ where δ is declination and α is right ascension) of magnitude −0.54 (+−0.14) mK. Observations of second and higher order harmonics are particularly exciting because, unlike the cosine anisotropy, they are not easily explained by local effects such as proper motion. Rather they give us information about the state of the Universe at very early times.

The existence of the cosmic microwave radiation was predicted by G. Gamow, R. Alpher and R. Hermann[5], and independently by R. Dicke, P. Peebles, P. Roll and D. Wilkinson[6] before the radiation was discovered by A. Penzias and R. Wilson[7] in 1965. Their theories have formed the basis of what is now referred to as the 'standard model' of the early Universe. In this model, the microwave radiation was once thermal radiation in equilibrium with the hot plasma that filled all space. About a half million years after the big bang, expansion of the Universe caused the previously opaque plasma to cool sufficiently to become transparent to the radiation, and since that time the signal has been travelling essentially unimpeded through the Universe. When one observes the radiation, one is literally looking at the shell of matter that last scattered it. The radiation comes from a more distant region of space, and was emitted earlier in time, than any other observed signal. The microwave radiation literally forms the spatial 'background' in front of which all other astrophysical objects lie.

There are many effects that could give rise to anisotropy in the radiation, including long-wavelength gravitational radiation, an overall rotation to the Universe, an anisotropic Hubble expansion and nonuniformity in the density of the Universe at the time of decoupling. Only the cosine or 'dipole' term (so called because a dipole distribution is proportional to the first spherical harmonic) can come from the Doppler shift due to the velocity of the observer with respect to the distant matter that last scattered the radiation.

Calculations indicate that the reported quadrupole anisotropy is consistent with that expected, based on observed inhomogeneities in the distributions of galaxies, and extrapolations in scale and in time to distributions at the time of decoupling[8,9]. In fact theorists were beginning to become uncomfortable with the absence of high harmonics in the anisotropy, since it was difficult to see how the present very lumpy Universe could have evolved from a highly uniform one in the short period of 10 to 20 billion years.

Our sigh of relief may be premature, however, because the experimental status is not totally resolved. Anisotropy measurements are very difficult, for they require measurement of millikelvin signals in the presence of system noise of typically several hundred Kelvin. Galactic emission and earth-shine are highly anisotropic, and difficult to eliminate. The history of anisotropy measurements is not reassuring, for published data have revealed large nonstatistical fluctuations subsequently shown to be due to systematic error. And most importantly, the present experiments do not agree within statistical errors. In 1979 G. Smoot and P. Lubin[10] at Berkeley published a measurement of Q_5 consistent with zero: +0.06 (+−0.2) mK. This value differs from the Princeton value by several standard deviations. Even the Princeton dipole terms differ significantly from those of Berkeley. No one doubts the reality of the dipole anisotropy, but the disagreement does show that one must be wary of accepting statistical errors as the true errors. The three experiments are very different in design, and the systematics could be very different. The Berkeley experiment sampled regions of the sky in both the Northern and Southern Hemispheres. The Princeton experiment had continuous sky coverage, but only in the northern sky. The Italian experiment had the poorest sky coverage of all; its high frequency makes it insensitive to synchrotron background, but more sensitive to thermal radiation from galactic dust. Since galactic emission can be point-like as well as diffuse, it is impossible to guess which experiment is most sensitive to unknown backgrounds. Smoot and Lubin are cautious, and willing to give as their

---

Richard A. Muller is Professor of Physics at the University of California, Berkeley and Faculty Senior Scientist at the Lawrence Berkeley Laboratory.

0028-0836/81/260609-02$01.00

final upper limit 1 mK for a quadrupole term. Fabbri *et al.* are similarly cautious. They consider their effect, 4.5 standard deviations from zero, to be only 'suggestive' of a quadrupole anisotropy. Only the Princeton group is convinced that they have been able to take all systematic errors into account.

If their quadrupole term is verified, then the Princeton group will deserve credit for their unambiguous announcement of its discovery. But I think it is premature to accept the effect as proven. Caution is called for, in part because the experiments are difficult, in part because of the discrepancies between the existing experim[illegible] and in part because the results are clos[illegible]hose expected. Fortunately we may no[illegible]e long to wait for verification. The Pri[illegible]on group has already flown in their ba[illegible]n gondola a maser receiver with consid[illegible]y improved sensitivity, and the Berkel[illegible]oup is about to fly a 90 GHz cryogen[illegible] system. The frequency behaviou[illegible]f the quadrupole term will clearly ide[illegible]ify it as cosmic background or as galactic interference, and in a few years we should obtain a detailed map of the 3K signal, including higher harmonics, from NASA's Cosmic Background Explorer satellite (COBE). The coverage and sensitivity of COBE are likely to give cosmologists a whole handful of new facts to inspire and constrain their theories. ☐

1. Smoot, G.F., Gorenstein, M.V. & Muller, R.A. *Phys. Rev. Lett.* **39**, 898 (1977).
2. Cheng, E.S., Saulson, P.R., Wilkinson, D.T. & Corey, B.E. *Astrophys. J. Lett.* **232**, L139 (1979).
3. Fabbri, R., Guidi, I., Melchiorri, F. & Natale, V. *Phys. Rev. Lett.* **44**, 1563 (1980).
4. Boughn, S.P., Cheng, E.S. & Wilkinson, D.T. *Astrophys. J. Lett.* **243**, L113 (1980).
5. Alpher, R.A., Bethe, H.A. & Gamow, G. *Phys. Rev.* **73**, 803 (1948); Alpher, R.A. & Herman, R.C. *Rev. Mod. Phys.* **22**, 153 (1950) and earlier references therein.
6. Dicke, R.H., Peebles, P.J.E., Roll, P.G. & Wilkinson, D.T. *Astrophys. J.* **142**, 414 (1965).
7. Penzias, A.A. & Wilson, M.L. *Astrophys. J.* **142**, 419 (1965).
8. Peebles, P.J.E. *Astrophys. J. Lett.* **243**, L119 (1981).
9. Silk, J. & Wilson, M.L. *Astrophys. J. Lett.* **244**, L37 (1981).
10. Smoot, G.F. & Lubin, P.M. *Astrophys. J. Lett.* **234**, L83 (1979).

Sexual differentiation of the brain

from Bruce S. McEwen

TESTOSTERONE acts during a limited period of perinatal development in male mammals to influence permanently the development of reproductive behaviour and neuroendocrine function, as well as to promote development of the reproductive tract. Although for some years the brain, and not the pituitary, has been presumed to be a site of testosterone action during development, it is only recently that there has been significant progress in understanding what actually happens to the brain during sexual differentiation.

Neuroanatomical studies have led the way, beginning with the first observation of sex differences in the sizes of nucleoli and cell nuclei in brains of gonadally intact[1] and hormone-treated[2] adult rodents and primates[3]. This initial work left open the question of whether hormone action during early development or during adult life was responsible for the morphological features. This question was answered in rat by the electron microscopic studies of Raisman and Field[4], which revealed that a sex difference in patterns of synaptic organization in the preoptic area of the adult rat depends, at least in part, on exposure to testosterone immediately after birth. Since then, studies at the light and electron microscopic levels have revealed sex differences in central nervous system morphology in songbirds[5,6] and rodents[7-12]. Among the more notable features of some of these sex differences is their occurrence in neural pathways and brain regions related to sexually dimorphic and hormone-dependent neural events such as song in birds[5,6], penile movement[11] and gonadotropin secretion[10] in rats.

How do these sex differences arise during the critical period of early perinatal neural development in mammals and birds? One mechanism being investigated is testosterone-stimulated neurite growth, which has been demonstrated in organ culture[13] where it appears to emanate from developing neurones that contain hormone receptors[14]. According to current ideas regarding the sequence of cellular events in neural development[15], neurite growth stimulated by gonadal hormones might be a sufficient condition for the production of differences in cell number and cell size as well as in the patterns of synaptic connections that are the major features of the sex differences described in the adult brain[4-12]. In spite of the appeal of such a scheme, it would be premature to overlook possible testosterone effects on other aspects of neural development, including the reinitiation of cell division or the promotion of cell death, and the production of substances, like nerve growth factor, which stabilize synaptic contacts. Furthermore, testosterone might alter the neurotransmitter phenotype of developing neurones — a possibility suggested by the actions of glucocorticoids on neurotransmitter type in developing sympathetic ganglia[16,17].

Hormonal stimulation of sexual differentiation occurs at a stage of neuronal development when considerable cellular differentiation has already occurred — oestrogen[18,19] and androgen[20,21] receptors and enzymes involved in testosterone metabolism to oestradiol[22] and 5-α-dihydrotestosterone[23] have already been laid down in certain neural cells. These enable particular neurones to respond to the hormonal signal, and it is known that both androgens and oestrogens arising from testosterone participate in sexual differentiation[24], albeit in differing degrees depending on the species[25]. Oestrogen and androgen receptors are already present in the hypothalamic area of female rodents in the last trimester of gestation[19,20] from around the time of final cell division[26]. There are also indications that oestrogen receptor levels in male and female rat brain start to increase shortly before birth[18] and that androgen receptor levels increase from one week after birth[21,27]. There are at present no indications for rat or mouse that, at these early ages, sex differences exist in levels of neural oestrogen and androgen receptors. Rather, it is the sex differences in testosterone levels in late fetal life[28] resulting in differential occupation of receptor sites in male and female brains[29,30] which provide the bias leading to sex differences.

Uncovering the factors, hormonal or otherwise, which govern the first appearance and subsequent perinatal increases of oestrogen and androgen receptors and testosterone-metabolizing enzymes in both male and female brains represents a major challenge for our further understanding of the origins of brain sex differences.

Burce S. McEwen is at the Rockefeller University, New York.

1. Pfaff, D.W. *J. Endocr.* **36**, 415 (1966).
2. Dörner, G. & Staudt, J. *Neuroendocrinology* **3**, 136 (1968).
3. Bubenik, G.A. & Brown, G.M. *Experientia* **29**, 619 (1973).
4. Raisman, G. & Field, P.M. *Brain Res.* **54**, 1 (1973).
5. Nottebohm, F. & Arnold, A.P. *Science* **194**, 211 (1976).
6. Gurney, M.E. & Konishi, M. *Science* **208**, 1380 (1980).
7. Greenough, W.T., Carter, C.S., Steerman, C. & DeVoogd, T. *Brain Res.* **126**, 63 (1977).
8. Gorski, R.A., Gordon, J.H., Shryne, J.E. & Southam, A.M. *Brain Res.* **148**, 333 (1978).
9. Rethelyi, M. *Neuroendocrinology* **28**, 82 (1979).
10. Matsumoto, A. & Arai, Y. *Brain Res.* **190**, 238 (1980).
11. Breedlove, S.M. & Arnold, A.P. *Science* **210**, 564 (1980).
12. Loy, R. & Milner, T.A. *Science* **208**, 1281 (1980).
13. Toran-Allerand, C.D. *Brain Res.* **106**, 407 (1976).
14. Toran-Allerand, C.D., Gerlach, J.L. & McEwen, B.S. *Brain Res.* **184**, 517 (1980).
15. Lund, R.D. *Development and Plasticity of the Brain* (Oxford University Press, New York, 1978).
16. McLennan, I.S., Hill, C.E. & Hendry, I.A. *Nature* **283**, 206 (1980).
17. Jonakait, G.M., Bohn, M.C. & Black, I.B. *Science* **210**, 551 (1980).
18. MacLusky, N.J., Lieberburg, I. & McEwen, B.S. *Brain Res.* **178**, 129 (1979).
19. Vito, C.V. & Fox, T.O. *Science* **204**, 517 (1979).
20. Vito, C.C., Wieland, S.J. & Fox, T.O. *Nature* **282**, 308 (1979).
21. Lieberburg, I., MacLusky, N.J. & McEwen, B.S. *Brain Res.* **196**, 125 (1980).
22. Reddy, V.V.R., Naftolin, F. & Ryan, K.J. *Endocrinology* **94**, 117 (1974).
23. Denef, C., Magnus, C. & McEwen, B.S. *Endocrinology* **94**, 1265 (1974).
24. Goy, R.W. & McEwen, B.S. (eds) *Sexual Differentiation of the Brain* (MIT Press, Cambridge, 1980).
25. Baum, M.J. *Neurosci. Biobehav. Rev.* **3**, 265 (1979).
26. Ifft, J.D. *J. comp. Neurol.* **144**, 193 (1972).
27. Attardi, B. & Ohno, S. *Endocrinology* **99**, 1279 (1976).
28. Weiss, J. & Ward, I.L. *Endocrinology* **106**, 306 (1980).
29. Westley, B.R. & Salaman, D.F. *Nature* **262**, 407 (1976).
30. Lieberburg, I., Krey, L.C. & McEwen, B.S. *Brain Res.* **178**, 207 (1979).

0028-0836/81/260610-01$01.00

Portuguese workshop on young stars

from a Correspondent

T TAURI stars are cool stars distinguished by spectra which show emission lines resembling those of the solar chromosphere and indicate a surrounding, dense, high-temperature region. They vary in light and are generally thought to be young because they are associated with dust clouds, lie above the main sequence of normal hydrogen-burning stars in the Hertzsprung–Russell diagram and show high abundances of lithium, an element destroyed by nuclear processes at and near the surface early in the main-sequence lifetime of the cool stars. T Tauri stars also show strong continua in X rays, the ultraviolet, infrared and radio which indicate a variety of processes in the circumstellar regions around the stars.

At a recent meeting* on T Tauri stars, one of the main questions discussed was whether the chromospheres contain inflowing or outflowing material. It might be expected that such newly formed stars would still be accreting interstellar material, yet in almost all stars the emission lines show a dip on the blue side. If this dip is due to absorption, it is perhaps natural to attribute it to outflowing material. However, Bertout (University of Heidelberg) and Krautter (European Southern Observatory) pointed out the great attractions of the infall model which provides a region of elevated temperature near the star where the infalling gas is brought to rest. Penston (Royal Greenwich Observatory) noted that all the MgII line profiles observed to date by the International Ultraviolet Explorer Satellite show an absorption or absence of emission on the blue wing and Lamers (University of Utrecht) agreed that in other types of star this is a symptom of mass loss. The most complete model presented for the flow around a T Tauri star was the stellar wind model presented by Lago (University of Oporto) for the star RU Lupi. The spectra of this star provide a series of constraints on a wind model linking the flow velocity (from the line widths) with the density and temperature (from various line ratios). A satisfactory model is obtained if the wind is driven by Alfvenic waves. Schwartz (Queen Mary College, London) remarked that the energy flux carried away in this wind as kinetic energy alone is very high, being some 20 times the luminosity of the Sun.

Observations from space are now making important contributions. Gahm (University of Stockholm) reported that the ultraviolet spectra of some stars showed evidence of a hot gaseous continuum and emission lines from gas at 100,000 to 200,000 K are seen in most stars. The situation for hotter gas at a million degrees (that is, a corona) was not so clear. A few stars were detected as X-ray sources by the Einstein Observatory but most were not. An absence of X-ray emission might be due to absorption in an overlying wind but attempts had also been made to find emission from coronae in forbidden lines from highly ionized species like Fe^{9+} and Fe^{13+}. These would not be absorbed in the same way and in some stars the limits were so good that one could say that the corona was much weaker relative to the chromosphere than it was in the Sun. A compatible result was presented by Kuhi (Lick Observatory) who showed from his analysis of the X-ray data that there was an anti-correlation of the X-ray luminosity of the detected T Tauri stars with Hα equivalent width. In fact the X-ray fluxes, and maybe therefore the coronae, are weaker than for normal stars of the same spectral type. These results suggested some important differences between the winds round T Tauri stars and the solar wind.

Concerning the presence of dust around T Tauri stars, the workshop accepted compelling evidence that, at least for some stars, alternative explanations for the polarization and infrared excesses were no longer tenable. Bastien (University of Bonn) reported that as the polarization was unchanged between Hα and the nearby continuum, the polarizing region must be outside the emission line zone and scattering by electrons cannot be the polarizing mechanism. Cohen (NASA — Ames) showed infrared spectra with strong absorption from ice and silicates which demonstrate that solid particles and not electrons are also the source of the infrared emission. In addition, it is difficult to account for the emission observed from aircraft at wavelengths near 100 μm in any other way, although there may still be a role for gaseous emission processes (free–free emission) at radio wavelengths. Panagia (University of Bologna) discussed this point, noting that for T Tauri, the best observed case with an $S_\nu \propto \nu^{0.7}$ spectrum, the data could be explained by free–free emission in a constant velocity wind. However, in other cases where the radio spectral index is as yet unknown, other models are possible and include the infall model or the origin of the radio emission in 10^9 K flares. Bertout and Cohen presented data indicating that some of the radio sources were variable. The energy necessary to maintain the ionized region can be very large; approximately the same amount as that emitted at the photosphere.

Lang (Tufts University) discussed the environment from which T Tauri stars are born. They lie in dense dark molecular clouds with densities of order $n_{\mathrm{H2}} \sim 10^3$ and temperatures near 20 K. The ultraviolet emission spectra also show evidence for molecular hydrogen. Brown (University of Oxford) advanced an explanation in which the strong Lyα radiation from the star fluorescently excites the molecules in the surrounding medium. Detailed investigation of the mechanism gives the excitation temperature in the molecular cloud and suggests shocks as the heating mechanisms. Cohen also discussed the relationship to T Tauri stars of the Herbig–Haro objects, small nebulae often found in the vicinity of the young stars. He reported that both previously proposed theories of these objects had been found correct in different cases. Some H-H nebulae were 'mirrors' scattering the light of the T Tauri stars; others were shock-excited regions. In the latter, the mechanical energy flux required to be emitted from the stars to drive the exciting shocks is very high and Cohen reported that Herbig had in one case found proper motions representing transverse velocities of 300 km s^{-1}. If the wind was spherical the energy carried away was extremely high!

Throughout the workshop, participants were comparing T Tauri stars with other types of star which shared some of their properties. Doazan and Thomas (University of Paris) stressed the similarities with Be stars (hot emission line stars) in which several of the same circumstellar phenomena are seen — indeed one subgroup, the Herbig Ae and Be stars, are almost certainly hotter analogues of the T Tauri stars. However, 'classical' Be stars seem to be distinguished by the absence of dust. Felli (Arcetri Observatory) also pointed out the similarity to BN objects. These are infrared stars without optical counterparts in H II regions and are named after the prototype in the Orion Nebula discovered by Becklin and Neugebauer.

Kuhi gave an account of his recent, very thorough programme on the rotational velocity of T Tauri stars. He noted that the oft-reported fact that T Tauri absorption lines are 'washed out' could either mean that they were broad or were filled in by emissions. He had avoided some of these types of difficulty by examining the Fourier transform of the absorption spectrum. He confirmed that some T Tauri stars had rotational velocities much higher than those normal for late-type stars but also found that many were rotating too slowly

---

*A workshop on T Tauri stars was held in mid-April at Ofir near Oporto, Portugal. The meeting was organized by the University of Oporto and jointly sponsored by the Portuguese government, two Portuguese agencies, INIC and JNICT, and the Fundacao Gulbenkian.

0028-0836/81/260611-02$01.00

(≤ 25 km s^{-1}) to be detected. Only those stars, more massive (according to their evolutionary tracks) than about 1.5 M , were likely to be rapid rotators — a result mirroring what occurs on the main sequence and strongly suggesting that stars have solved their angular momentum problems before the T Tauri phase. His result may suggest that rapid rotation is not the main cause of the T Tauri phenomenon but one should be cautious because the stars with the strongest emission effects do not have photospheric absorption lines enabling their rotation rates to be measured.

Many lines of evidence suggest that the main distinguishing feature of T Tauri stars is the copious flux of mechanical energy — winds, waves and shocks — which emanate from them. In some cases this is known to be a substantial fraction (10 per cent or more) of the radiative flux passing the photosphere. With this extraordinary result well established, it is perhaps surprising that the phenomenon has not received more serious theoretical attention — what possible processes can lead to such a large proportion of the energy losses in mechanical forms? With Kuhi's important work perhaps hinting that rotation is not the primary cause of the effect, is it magnetic fields to which the theoreticians must now look for the explanation?

Participants stressed the desirability of coordinated observations at different wavelengths. In particular, the importance of space observations and the pressure of time on such facilities — like IUE and EXOSAT — made this vital. Possibilities for future coordinated observing programmes were discussed and it was concluded that May 1982 would provide the first major opportunity to deploy IUE and other facilities together and would be particularly suited for Southern stars. It was agreed that Kuhi of the Lick Observatory at Santa Cruz, California would be central coordinator and that any other interested parties would be welcome to join in coordinated campaigns.

Biogeography: in search of principles

from Colin Patterson

'BIOGEOGRAPHY is a strange discipline. In general, there are no institutes of biogeography; there are no departments of it. There are no professional biogeographers — no professors of it, no curators of it. It seems to have few traditions.' (see ref. 1). Biogeography in the accepted sense was invented just over a century ago, by Alfred Russel Wallace, who saw his *Geographical Distribution of Animals* (1876) as elaborating the two chapters on geographical distribution in *The Origin of Species*, just as Darwin had elaborated chapter 1 in *The Variation of Animals and Plants Under Domestication*. If biogeography has traditions, they are rooted in Wallace's works. As all traditions sooner or later must, these have come under attack during the past 15 years.

Wallace's programme was based in the geologist Dana's theory of fixed continents and permanent ocean basins, an idea endorsed by Darwin. If the continents and oceans have always been roughly where they are, and if 'every species has come into existence coincident both in space and time with a pre-existing closely-allied species' (Wallace's 1855 formulation of evolution), the biogeographer's task is to infer when and where each species or group arose, and to explain the subsequent distribution of each. Help was sought in the fossil record, in anecdotal and historical records of dispersal, and in empirical investigation such as Darwin's immersion of seeds in salt water, and his scraping potential passengers from ducks' feet. The results of this programme became embodied in general theories: for example, competition and natural selection are more effective on large continental areas such as the northern landmasses (Darwin's 'more efficient workshops of the north'); rigorous climates promote evolution (the monoboreal hypothesis — life has radiated from the Arctic). According to these theories, the close relationship between the southern temperature biotas of Australasia, South America and Africa, evident to the botanist Joseph Hooker in the 1850s, was illusory; those regions were culs-de-sac where the oldest groups survived, forced southwards by the newer and more efficient products of northern workshops. These ideas, or variants of them, were dominant in biogeography until the middle 1960s, when geophysical evidence of continental mobility became overwhelming. With that came the realisation that the relationship between southern biotas reflected Earth history.

In the ensuing shakeout, three schools of biogeographical theory have emerged[2]. The first is an updated version of Wallace's programme, centres of origin and mobile life, married with the mobile Earth of plate tectonic theory. The second, equilibrium theory, has grown from MacArthur and Wilson's 1967 book, *The Theory of Island Biogeography*. Equilibrium theorists apply the mathematical models of population genetics and ecology to biogeography, and investigate the chances of dispersal, colonisation and extinction empirically, through work on natural or artificial islands. The third school, vicariance biogeography, finds its roots in the voluminous and idiosyncratic prose of Leon Croizat[3], one of whose slogans is 'Earth and life have evolved together'. According to vicariists, the distribution of life is governed not by the wanderings of animals and plants, but by Earth history. Tectonic events have split whole biotas, and historical relationships between areas may be reconstructed through the systematic relationships of species. Vicariance theory and method were reviewed at a major meeting in New York in 1979[4].

A recent meeting* at the British Museum, set against this background, had spokesmen for all three schools. It did not develop into a showdown or conflict between the three; rather, speakers stuck close to their preferred school, with nods or sallies towards the others. Many speakers commented on the three approaches and the conflict between them, but no general argument or consensus emerged.

In view of the ideological fragmentation of the meeting, I mention only a few of the papers. Apologies are due to many speakers whose contributions are not mentioned, but some of them, botanists in particular, were still in the data-gathering rather than explanatory phase.

Amongst equilibrium theorists, E.C. Pielou (University of Dalhousie) contributed to the debate on punctuational change versus gradualism. She provided a method of analysis — ordered similarity matrices — which avoids qualitative judgments, and applied the method to forams from the Bay of Fundy. Step-like rather than gradual change was inferred in these forams, thought to be due to sudden change in tidal pattern. Similar results were reported by the palaeontologist R. Coope (Birmingham University), summarising his work on Quaternary and Pleistocene Palaearctic beetles. These animals show no detectable evolutionary change or extinction; instead the effect of climatic change is paramount. Coope believes his insect communities moved to dodge environmental change, rather than evolving or becoming extinct. D. Simberloff (Florida State University) presented a

---

*As part of the British Museum (Natural History) centenary celebrations, the Museum joined with the Systematics Association in organising a week-long international symposium on biogeography, 6-10 April, under the title 'Time and Space in the Emergence of the Biosphere.' About 100 biologists and geologists met to hear almost 30 speakers.

Colin Patterson is in the Department of Palaeontology, British Museum (Natural History), London.

0028-0836/81/260612-02$01.00

pessimistic summary of equilibrium theory, proposing that the equilibrium model is hard to define and difficult to test, but seems to have failed some tests. Simberloff was concerned with hypothesis testing, and the difficulty of using patterns to determine causes, themes which came up repeatedly.

I.R. Ball (University of Amsterdam) addressed the most general problems of biogeography and tried to steer a middle course between the shoals of various schools. Like Simberloff, Ball discussed hypothesis testing and the problem of distinguishing pattern from process, but he also spoke of recent emphasis (overemphasis in his view) on Method. By Method, Ball meant the cladistic debate, which was mentioned as often at the meeting as in recent issues of *Nature*. The impact of cladistics on biogeography has been in two stages. The first, best exemplified in Brundin's 1966 monograph on austral chironomid midges[5], used cladograms of taxa to estimate their centres of origin and routes of dispersal. In brief, the method is to correlate the cladogram and map and to use a parsimony procedure, minimising the number of dispersals, to localise branch points in the cladogram. The method is infallible, but is unduly sensitive to fossil localities and depends on the axiom that each taxon had a centre of origin from which it dispersed, by means which are assumed to depend in part on the capacities of the particular organisms, and in part on chance. For example, by this method man is inferred to have originated in and dispersed from Africa, since our cladistic relatives, chimpanzees and gorillas, are found there, as are the most primitive man-like fossils.

The second stage of cladistic biogeography was developed within vicariance theory. This method seeks to combine cladograms of individual taxa into cladograms of areas[6]. In formal terms, cladistic analysis moves to a more general level, from that at which groups of organisms are the units, ordered by finding congruence between character cladograms (cladistic systematics), to one at which geographical areas of endemism are the units, ordered by finding congruence between cladograms of organisms inhabiting those areas. At this level of analysis, a cladogram of apes and humans would be of interest, not as a source of inferences about the areas of origin and routes of dispersal of those animals, but as one among many taxon cladograms bearing on the interrelationships of the areas occupied by apes and humans. In this mode of analysis, centres of origin have the same status as ancestral species in cladistic systematics: they are regarded as irrelevant, because ancestors are components of trees, not of cladograms, and there are many trees compatible with one cladogram.

The problem with the application and testing of the method just described is shortage of the raw material, taxon cladograms. At the meeting, taxon cladograms were presented by several speakers, including J.R.G. Turner (Leeds University), Amazonian butterflies; J.D. Holloway (Commonwealth Institute of Entomology), Indo-Australian butterflies; and A.R. Milner (Birkbeck College, London), who found congruence between his cladogram of urodele amphibians and inferred geological history of the northern continents. But the only attempt to combine cladograms of different taxa was by C.J. Humphries (BM(NH)), who took the southern beeches (*Nothofagus*) as his starting point, and arrived at a general cladogram of the southern continents. K. Thomsen (Aarhus University) applied the cladistic method and Rosa's theory of hologenesis to *Homo sapiens*, inferring a cosmopolitan ancestral population which was subdivided by successive vicariance events. This analysis engendered some heat amongst the anthropologists present. Our own species was also discussed by G. Nelson (American Museum of Natural History), who presented a historical survey of biogeographical thought from Linnaeus to the present, and used theories of the distribution of *Homo sapiens* to illustrate his theme that centre-of-origin biogeography is akin to Linnaeus's ideas on the location of Paradise.

My overall impression of the meeting was of a discipline in search of some discipline, or some principles. At least one speaker, Nelson, held out hope for the future, remarking that 'biogeography is still in its beginning phase, one comparable to astronomy at the time of Galileo'. To Nelson, the hope lies in area cladograms. To others, such as I.R. Ball, the hope is in a pluralistic approach. □

---

1. Nelson, G. *J. hist. Biol.* **11**, 269 (1978).
2. Vuilleumier, F. *C. r. somm. Séanc. Soc. Biogéogr.* **54**, 41 (1978).
3. Croizat, L. *Space, Time, Form: The Biological Synthesis* (Caracas, 1964).
4. Nelson, G. & Rosen, D.E. *Vicariance Biogeography: A Critique* (Columbia University Press, New York, 1981).
5. Brundin, L. *K. svenska VetenskAkad. Handl.* (4) **11**, (1966).
6. Nelson, G. & Platnick, N.I. *Systematics and Biogeography: Cladistics and Vicariance* (Columbia University Press, New York, 1981).

NATURE

100 years ago

THE GORILLA AND THE CHIMPANZEE

Mr. H. von Koppenfels, who is now engaged in explorations in the Gaboon district of Western Africa, in a letter published in the last number of the *American Naturalist,* states that he has good evidence of the existence of crosses between the male gorilla and the female chimpanzee. "This," he says, "settles all the questions about the gorilla, chimpanzee, Kooloo-Kmba, N'schigo, M'bouvé, the Sokos, Baboots, &c." Herr v. Koppenfels observes that the "French *savants* seem to have a special predilection for creating new species from variations in the form of the skull, such as often occur in this group of animals. The chimpanzee is found all over Tropical Africa, and naturally exhibits considerable variation. Du Chaillu's Kooloo-Kamba, N'schigo, and M'bouvé are *not* distinct species, and this traveller, who is certainly a man of merit, but is too credulous, has been imposed upon by the mendacity of the natives, which beggars description. The mongrel progeny of the male gorilla and female chimpanzee discovered by me is found but in individual cases, and as such deserves no special name."

from *Nature* **24**, 30 June, 203, 1881.

Planet X: is it necessary?

from David W. Hughes

EPHEMERIDES are tables recording the predicted positions of celestial objects against the sky background, and astronomers are still making strenuous efforts to improve their accuracy. Careful comparisons between, for example, the observed positions of Neptune and Uranus and the predicted positions of these planets were a prelude to the discovery of Pluto in 1930. It is thought by some astronomers that a continuation of this process, including Pluto, could now lead to the discovery of a 'Planet X', the next planet out. Others doubt the existence of more planets and believe that the tally of major planets is complete.

David W. Hughes is in the Department of Physics, University of Sheffield.

The last orbit of Pluto was published in 1972 and was obtained by analysing the positions of the planet recorded between 1914 and 1965. Over 500 observations were used, including 14 prediscovery ones where Pluto had been photographed but had not been recognized as a planet.

A new set of orbital parameters for Pluto has been computed by P.K. Seidelmann, G.H. Kaplan, K.F. Pulkkiner, E.J. Santoro and T.C. Van Flandern of the United States Naval Observatory, Washington DC (*Icarus* **44**; 19, 1980). Another 15 years of observations have been added to the previous set, the majority coming from a Naval Observatory programme for observing Pluto, the same programme that fortuitously lead to the discovery of

0028-0836/81/260613-02$01.00

Pluto's satellite. This discovery and the analysis of the satellite's orbit removed a previously unknown quantity, the mass of Pluto, from the orbital calculation.

Pluto has now been observed over about a quarter of its orbit. The new results give the orbit a semi-major axis and period of 39.718 AU and 250.3 yr whereas Allen's *Astrophysical Quantities* (1973) quotes 39.44 AU and 247.7 yr respectively. The inclination is 17.139029°, the argument of perihelion 112.985694° and the eccentricity 0.25234645°, all for the epoch 1979 May 7.

Even now the residuals between the recent observations and the new ephemeris indicate that improvements must be made and to this end positional observations of the planet must continue.

Prediscovery estimates stated that Pluto's mass would be about five times that of the Earth. It is now known to be about a five-hundredth of the mass of the Earth. The ephemerides of Uranus and Neptune have been computed under the assumption that the mass of Pluto is very small. Comparison in right ascension and declination have also been made between the observed positions of the two planets and the ephemerides. The results can be interpreted as indicating that a secular difference of the orders of half an arc second exists between observation and ephemeris for the orbits of both Uranus and Neptune.

The authors conclude that some unmodelled force is acting on the outer planets. Observations of six long-period comets, which have had more than one apparition, show that these too are being perturbed by an unmodelled force perpendicular to their orbital plane. The well known non-gravitational force acting on comets has no component in this direction. All these comets have aphelia beyond Neptune.

The cause of the force is unknown. That a Planet X beyond Pluto is exerting a gravitational effect is a possibility but the authors stress that the long and unsuccessful photographic search made by Clyde Tombaugh makes its existence unlikely. If, for example, Planet X was as massive as Jupiter, Tombaugh concluded that he would have found it if it were closer than 470 AU. Also, this phantom 'Planet X' does not seem to have a measurable effect on the orbits of new comets. These comets, which have just been 'bumped' out of the Oort cloud by a passing star, seem to have near perfect parabolic orbits and there is no anisotropy in their major axis distribution. Furthermore, Planet X does not seem to affect the plane of symmetry of the Solar System. This seems to be dominated by Jupiter.

Our understanding of the motions of Uranus, Neptune and Pluto remain incomplete but the authors remind us that we should not leap to a 'Planet X' hypothesis too quickly. Leverrier made this error when, in 1859, he suggested that the anomalous advance of the perihelion of Mercury was due to perturbations exerted by a planet or planets inside Mercury. Even after the excitement caused by Lescarbault's supposed discovery of Vulcan, the correct explanation had to await Einstein's theory of relativity. □

Muscle regulation: a decade of the steric blocking model

from John Squire

THE first insights into the molecular mechanism of muscle regulation were obtained nearly ten years ago. The calcium concentration in the sarcoplasm, itself regulated by the nervous input to the muscle via the T-tubules and the sarcoplasmic reticulum, was known to be involved through the concentration-dependent binding of calcium to the actin filament protein troponin[1]. However, little was known about how troponin influenced the structure of these filaments.

Evidence existed that troponin is associated with the rod-shaped α-helical protein tropomyosin[1], and that in the actin filament the actin subunits form a structure which approximates to two coaxial long-pitched helical strands of globular beads[2] (Fig.1). The tropomyosin molecules, each with a single troponin complex attached, were thought to run head-to-tail to form continuous protein strands along the two long-pitched grooves of the actin helix, such that each single tropomyosin/-troponin repeat would be associated with seven actin subunits along each helical strand. But how could calcium binding to a single troponin molecule influence the binding of myosin heads to seven actin subunits? Clearly tropomyosin could be involved, but in what way?

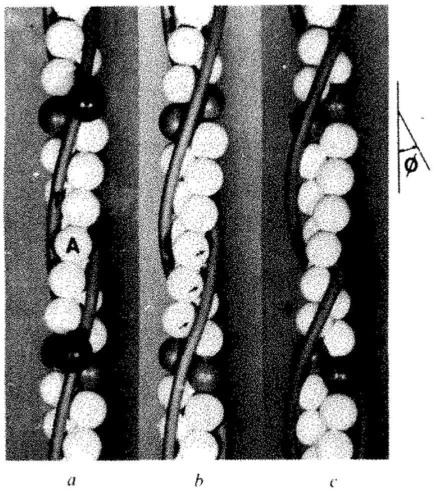


Fig.1 *a*, Early model of the thin filament; actin monomers (A) form a helical structure[2] with tropomyosin strands (T) running centrally up the two long-period grooves of the helix[1]. Troponin molecules (dark spheres) are attached to each tropomyosin molecule and there is one tropomyosin/troponin pair for every seven actin monomers. Grey actins mark the seven monomer repeat. *b* and *c*, Revised models[4,5,7] in which the tropomyosin is not centrally placed but is at about $\theta = 70°$ (*b*; see Fig. 2*a*), and about $\theta = 50°$ (*c*). In the steric blocking model the tropomyosin position in *b* is to one side of the sites (arrows) where myosin heads can bind to actin; in *c*, these sites are blocked by tropomyosin. Troponin is not shown in *b* and *c*. According to Seymour and O'Brien[13] the Z-line would be at the lower end of these filaments.

Part of the answer came when it was found that differences in the X-ray diffraction patterns of relaxed and contracting muscles[3-6] could only be satisfactorily interpreted[4,5,7] in terms of a change in position of tropomyosin in the grooves of the actin helix. Figures 1*b*, *c* and 2*a* illustrate the kind of movement involved. In relaxed muscle the tropomyosin (viewed end on in Fig.2*a*) was thought to lie well out of the groove ($\theta = 45 - 55°$), whereas in active muscle it moved much closer to the groove ($\theta = 65 - 75°$). Later studies[8,9] confirmed this tropomyosin movement, but how was it linked with thin filament regulation?

In 1970, Moore, Huxley and DeRosier[10] obtained the first three-dimensional reconstructions from electron micrographs of negatively stained actin filaments decorated with the S-1 head sub-fragment of myosin. Their reconstructions could be interpreted to mean that the myosin heads attach to actin as shown in Fig.2*b*. The observed tropomyosin shift was soon combined with this postulated myosin attachment position to produce the so-called steric blocking model of thin filament regulation[4,5,7]. The 'off' position for tropomyosin in Fig.2*a* is very close to the attachment site of myosin to actin in Fig.2*b*, but the 'on' position is clear of this site. Thus tropomyosin could physically block the attachment of myosin to actin in

0028-0836/81/260614-02$01.00

relaxed muscle (low calcium) but would roll into the groove and out of the way of the attachment site, due to calcium binding to troponin, when the muscle was stimulated. Although it was not known at the time that the myosin site and the tropomyosin strands were on the same side of the actin groove, this seemed a reasonable suggestion; otherwise such a steric blocking mechanism would be harder to envisage.

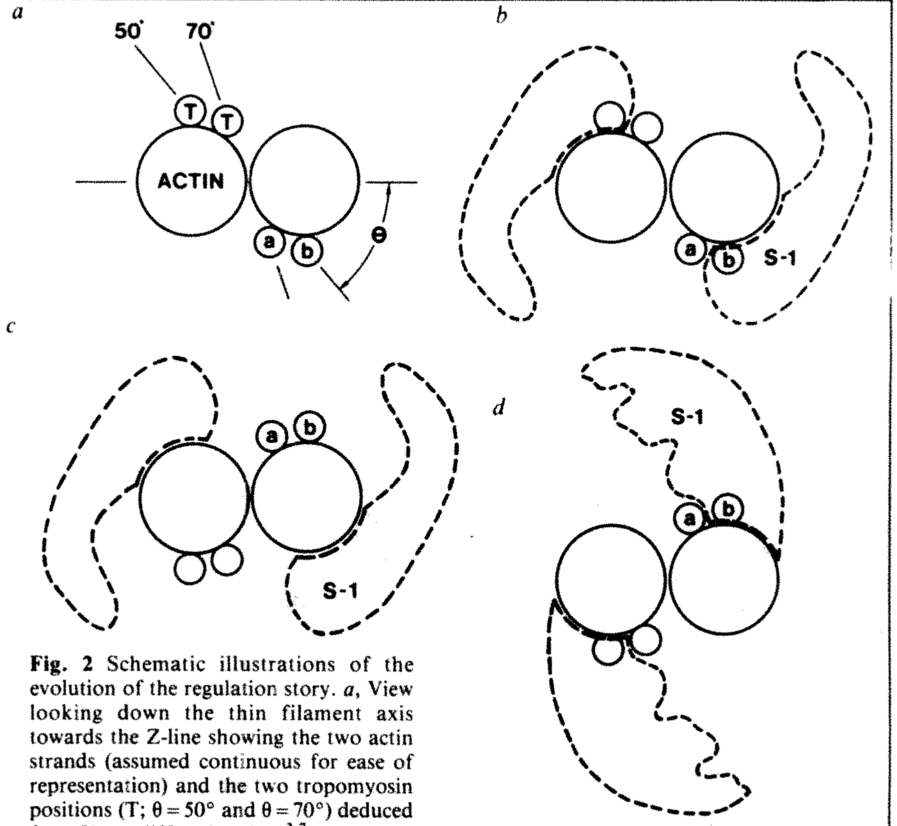


Fig. 2 Schematic illustrations of the evolution of the regulation story. *a*, View looking down the thin filament axis towards the Z-line showing the two actin strands (assumed continuous for ease of representation) and the two tropomyosin positions (T; $\theta = 50°$ and $\theta = 70°$) deduced from X-ray diffraction data[3-7]. Note that such data also allow the tropomyosin positions to be bottom left and top right and not as shown. *b*, The myosin head (S-1) attachment position according to Moore *et al.*[10] is close to tropomyosin position 'b' (the blocking position; $\theta = 50°$), but it is away from position 'a' (the active position; $\theta = 70°$). However, Seymour and O'Brien[13] showed the tropomyosin position to be as in *c*; it is across the thin filament groove from Moore *et al.*'s myosin head position. *d*, The latest reconstructions of decorated thin filaments by Taylor and Amos[14] now put the myosin head also across the groove, so that steric blocking can still occur. (Figures not drawn exactly to scale.)

Nobody assumed that steric blocking would prove to be the whole story and biochemical evidence was soon obtained which suggested that additional aspects of thin filament regulation might be mediated through the actin subunits themselves, possibly as a result of tropomyosin movement[11,12]. But it was not until 1980 that the steric blocking model received its first structural jolt. Seymour and O'Brien[13], from studies of thin filament structure in I-segments, showed that tropomyosin does indeed lie on the opposite side of the actin groove to the myosin-binding site suggested by Moore *et al.*[10] (Fig.2*c*). How could steric blocking occur now? Perhaps the biochemists were right and tropomyosin movement, about which there was little doubt, only acted indirectly through the actin subunits in modifying the binding site.

One of the inherent problems with three-dimensional reconstructions of helical structures such as actin filaments is that it is difficult to obtain reliable information about densities close to the axis of the helix. This is because, in the diffraction patterns used for the analysis, such information is at the outer edges (high radius positions along layer lines) where the effects of disorder in the specimen are most serious. For this reason Taylor and Amos[14] have recently repeated the work of Moore's group[10] using longer straight stretches of decorated thin filaments and more up-to-date methods, including minimal-dose electron microscopy. This enabled them to retain more information about the low-radius features in the decorated filaments — features relating to the actin subunits themselves — and they found that the peaks chosen to be actin by Moore *et al.*[10] were probably part of the myosin head and that new density features close to the thin filament axis (Fig.2*d*) could more sensibly be assigned to actin.

These results clearly provide a new insight into the mode of attachment of the myosin head to actin. As early as 1965, Reedy, Holmes and Tregear[15] suggested from electron micrograph appearances that in rigor muscle the long axis of the actin-bound myosin heads makes an angle (ϕ in Fig. 1) of about 45° to the filament axis. Combined with an apparent tilt closer to 90° in fixed relaxed muscle this result led eventually to the swinging crossbridge model of force generation in muscle[16]. However, modelling of X-ray diffraction patterns of rigor muscles[17,18] has generally tended to suggest an attachment angle (ϕ) rather closer to 90° than the 45° value suggested by Reedy *et al.*[15]. The shape of the myosin head itself makes the definition of tilt rather ambiguous; the head is probably curved, about 60 Å across proximal to the actin-binding site, tapering to < 30 Å at its distal end, and between 120 and 180 Å long[14,19]. But the Taylor and Amos results suggest that ϕ is about 55–65° and a similar study[20] puts the angle closer to 75°. Although at first sight these values conflict with the appearances in Reedy's micrographs of rigor insect muscle[15,21], it should be remembered both that the S-2 link between the head and the myosin filament backbone is present in those sections and that until such micrographs are analysed in three dimensions only two-dimensional projections of the structure are being studied. In such views, the apparent tilt angle (ϕ) of myosin heads which do not have their long axes exactly parallel to the plane of sectioning will appear to be less (closer to 45°) than it is in three-dimensional reconstructions of decorated filaments where the angle can be measured more directly. Thus there may be no real conflict between these various results.

One of the most significant features of the recent Taylor and Amos work is that the myosin attachment site (Fig.2*d*), like the tropomyosin position, has now also moved across the actin groove from the early model (Fig.2*a*). Tropomyosin and the myosin attachment site are close together once again and, although some aspects of regulation are clearly mediated by actin[11,12], the steric blocking model lives on. □

John Squire is a Lecturer in Materials Science and leader of the Biopolymer Group at Imperial College, London.

1. Ebashi, S. & Endo, M. *Prog. Biophys. molec. Biol.* **18,** 123 (1968).
2. Hanson, J. & Lowy, J. *J. molec. Biol.* **6,** 46 (1963).
3. Huxley, H. E. *Biochem. J.* **125,** 85 (1971).
4. Huxley, H. E. *Cold Spring Harb. Symp. quant. Biol.* **37,** 361 (1972).
5. Haselgrove, J. C. *Cold Spring Harb. Symp. quant. Biol.* **37,** 341 (1972).
6. Vibert, P. J., Haselgrove, J. C., Lowy, J. & Poulsen, F. R. *J. molec. Biol.* **71,** 757 (1972).
7. Parry, D. A. D. & Squire, J. M. *J. molec. Biol.* **75,** 33 (1973).
8. Spudich, J. A., Huxley, H. E. & Finch, J. T. *J. molec. Biol.* **72,** 619 (1972).
9. Wakabayashi, T., Huxley, H. E., Amos, L. A. & King, A. *J. molec. Biol.* **93,** 477 (1975).
10. Moore, P. B., Huxley, H. E. & DeRosier, D. J. *J. molec. Biol.* **50,** 279 (1970).
11. Bremel, R. D., Murray, J. M. & Weber, A. *Cold Spring Harb. Symp. quant. Biol.* **37,** 267 (1972).
12. Weber, A. & Murray, J. M. *Physiol. Rev.* **53,** 612 (1973).
13. Seymour, J. & O'Brein, E. J. *Nature* **283,** 680 (1980).
14. Taylor, K. A. & Amos, L. A. *J. molec. Biol.* **147,** 297 (1981).
15. Reedy, M. K., Holmes, K. C. & Tregear, R. T. *Nature* **207,** 1276 (1965).
16. Huxley, H. E. *Science* **164,** 1356 (1969).
17. Miller, A. & Tregear, R. T. *J. molec. Biol.* **70,** 85 (1972).
18. Holmes, K. C., Tregear, R. T. & Barrington-Leigh, J. *Proc. R. Soc.* **B207,** 13 (1980).
19. Elliott, A. & Offer, G. *J. molec. Biol.* **123,** 505 (1978).
20. Toyoshima, C. & Wakabayashi, T. *J. Biochem.* **86,** 1887 (1980).
21. Reedy, M. K. *J. molec. Biol.* **31,** 155 (1968).

0028-0836/81/260615-01$01.00

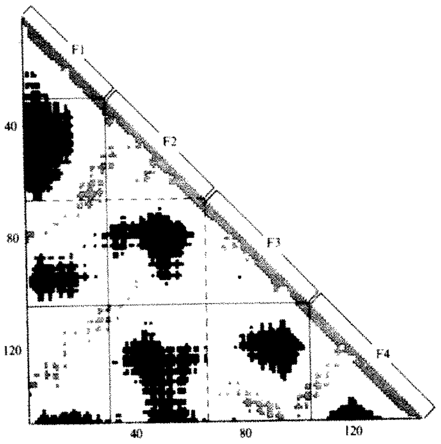
Exons and the structure, function and evolution of haemoglobin

from C.C.F. Blake

OVER the last few weeks there has been a flurry of papers in *Nature* that bear directly on the meaning of coding sequences or exons in the haemoglobin genes. In the mouse α- and β- globin genes there are three exons corresponding to amino acid residues 1–31, 32–99 and 100–141; and 1–30, 31–104 and 105–146, respectively. When the amino acid sequences of the α- and β-globins are aligned to maximize their structural homologies, their intron/exon junctions coincide. This immediately suggests that the 'meaning' of the exons is related in some way to the structure and function of the protein. The recent papers address themselves to the problem by analysing the exon-encoded regions in terms of the structure, function and evolution of the haemoglobin molecule.

Gō[1] has ingeniously used the diagonal plot (see the figure) to define regions of the polypeptide chain that are distant (> 27Å) from one another in the globin fold. He has found that there are four such regions, which are not domains, but which can be called sub-domains, or perhaps 'compact structures'. When included on the diagonal plot the exon boundaries neatly divide off these compact structures from one another, with the exception of the large central exonic region which is composed of two compact structures. Because of this Gō suggested that the central exonic region might consist of two 'fused' exonic regions with a division somewhere between residues 66 and 71. Quite remarkably this prediction has been rapidly verified by Marcker and colleagues[2] (see this issue of *Nature*, p.677) in their determination of the structure of the leghaemoglobin gene from soybean. This gene is composed of four exons, corresponding to amino acid residues 1–32, 32–68, 69–103, 104–C terminus. Marcker's results suggest that exons are generally very stable, but that they can undergo fusion or separation. Gō's successful prediction of an exon on the basis of protein structure adds powerful additional support to the idea[3] that they correspond to compact protein structures.

As Gilbert's suggestion[4] of a relation between exons and protein functional units appears to have been validated in lysozyme[5], consideration of this aspect of haemoglobin is of particular interest. In haemoglobin the functional correlation began with the suggestion[6] that the central exonic region appeared to correspond to a haem-binding unit. This suggestion was analysed in detail by Eaton[7], who not only confirmed this property of the central exon, but also showed that the distribution of intersubunit contacts showed a strong correlation with the exonic regions. Nearly all the $\alpha_1\beta_2$ contacts are in the central exonic region, and nearly all the $\alpha_1\beta_1$ contacts are in the third exonic region. This seems to suggest that while the contacts responsible for the early cooperativity of $\alpha_1\beta_2$ dimers may have developed within the haem-binding structure itself, the later tetramer cooperativity required the presence of the third coding sequence, which Eaton speculates may have replaced its predecessor by recombination. At the same time Beychok and his colleagues[8] began an experimental examination of the properties of the central exonic region of human haemoglobin, which can be excised from the globin chain by clostripain treatment. Their studies first demonstrated that the separated central region was indeed capable of binding haem tightly and specifically. An extension[9] of this analysis has now shown that although the central exon peptide can bind haem, it cannot stabilize the haem–dioxygen complex, which requires both the presence of the side exon products and the complementary subunit. The regainment of activity is accompanied by an increase in helical content, suggesting that a structural conformational change may be involved. With the verification of the haem-binding function of the central exonic region, it is interesting to consider the significance of its division into two in leghaemoglobin. Despite Marcker's assertion that it has no obvious functional correlation, the division in fact neatly segregates the proximal and distal haem contacts. This suggests that the divided central exonic region may be a primitive form and that the two exons fused along the haemoglobin line to ensure a single protein product capable of specifically interacting with haem.



A plot of all the distances between α-carbon atoms in the haemoglobin β chain. Both ordinate and abscissa are the residue sequence number. Distances between pairs of atoms greater that 27 Å appear as dark regions. Solid lines are drawn at the joints between polypeptide segments (F1, F2 + F3, and F4 respectively) encoded by different exons — between residues 30 and 31, and 104 and 105. Note that these lines scarcely cross the dark regions.

In a consideration of their results, Beychok and his colleagues note that the central exon product, isolated or in noncovalent association with the side peptide fragments, has spectral characteristics and other behaviour reminiscent of the b-type cytochromes. In 1975 Rossmann and Argos[10] demonstrated a structural similarity between the globins and cytochrome b_5, which was later extended to include cytochrome c_{551}[11]. In each case the closest agreement involved the haem-binding region and corresponds rather precisely to the central exon in the globin chain. When the protein folds were maximally superposed, the haem irons were found to be separated by 4.2Å in the globin – cytochrome b_5 comparison and by 5.3Å in the globin – cytochrome c_{551} comparison, and the haem normals by 13° and 9° respectively. The results of these chemical and structural comparisons suggest that the globins and the cytochromes may have evolved from a common haem-binding domain encoded by one or more exons, which by combining with other gene elements has generated a multiplicity of haem-binding proteins of diverse function. These conclusions are very similar to those obtained from the examination of the exon structure–function relationships in hen and T4 lysozymes[5], and lend strong support to Gilbert's proposal that in eukaryotes exons code for functional protein units which can serve in rapid protein evolution. The great puzzle now is why prokaryotic genes do not have this apparently advantageous structure. □

C.C.F. Blake is in the Laboratory of Molecular Biophysics, Department of Zoology, Oxford.

1. Gō, M. *Nature* **291**, 90 (1981).
2. Jensen, E.O., Paludan, K., Hyldig-Nielsen, J.J., Jorgensen, P. & Marcker, K.A. *Nature* **291**, 677 (1981).
3. Blake, C.C.F. *Nature* **273**, 267 (1978).
4. Gilbert, W. *Nature* **271**, 501 (1978).
5. Artymiuk, P.J., Blake, C.C.F. & Sippel, A.E. *Nature* **290**, 287 (1981)
6. Blake, C.C.F. *Nature* **277**, 598 (1979).
7. Eaton, W.A. *Nature* **284**, 183 (1980).
8. Craik, C.S., Buchman, S.R. & Beychok, S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1384 (1980).
9. Craik, C.S., Buchman, S.R. & Beychok, S. *Nature* **291**, 87 (1981).
10. Rossmann, M.G. & Argos, P. *J. biol. Chem.* **250**, 7525 (1975).
11. Argos, P. & Rossmann, M.G. *Biochemistry* **18**, 4951 (1979).

0028-0836/81/260616-01$01.00

REVIEW ARTICLE

High-speed photography

George H. Lunn
Association for High-Speed Photography, 57 Whitedown Road, Basingstoke RG26 6BZ, UK

The most powerful of human senses is sight, to such an extent that to indicate we understand even an abstract concept we say "I can see that!". Until photography became a generally available technique, it required considerable skill and practice to produce any form of picture or illustration. The combination of variable time and pictures into one subject makes one appreciate the fascination of high-speed photography. Most high-speed photographers are users; they apply the techniques supplied to answer questions in some other study.

PHOTOGRAPHY began with Niepce and Fox-Talbot in the 1830s; indeed Fox-Talbot, by using an electric spark in his famous Royal Institute lecture (1851) simultaneously introduced high-speed photography. However, until 1879, when Ilford produced an emulsion coating machine, and 1880, when Kodak produced dry plates, photography was restricted to the few who could be bothered to work with wet plates. The first to capture public attention was Muybridge with his classical studies of animal motion, although Jansen in 1874 had studied the transit of Venus with a revolver camera. This was the period of cine-camera development (Edison, Lumiere, among others) with Marey and Bull in Paris creating mechanical cameras capable of 700 pictures per second (p.p.s.) (1894) and up to 2,000 p.p.s. (1904) with Worthington setting the pace in spark-photography studying splashes.

The early high-speed cameras used rotating disks carrying either lenses or film with the alternate one stationary to produce good short sequences such as the flight of a fly. Boys produced a rotating disk-bearing lens which much later was used by Schonland to show how a lightning strike behaved. These were limited in the number of pictures and top speed and restricted at that time by the low sensitivity of available photographic materials.

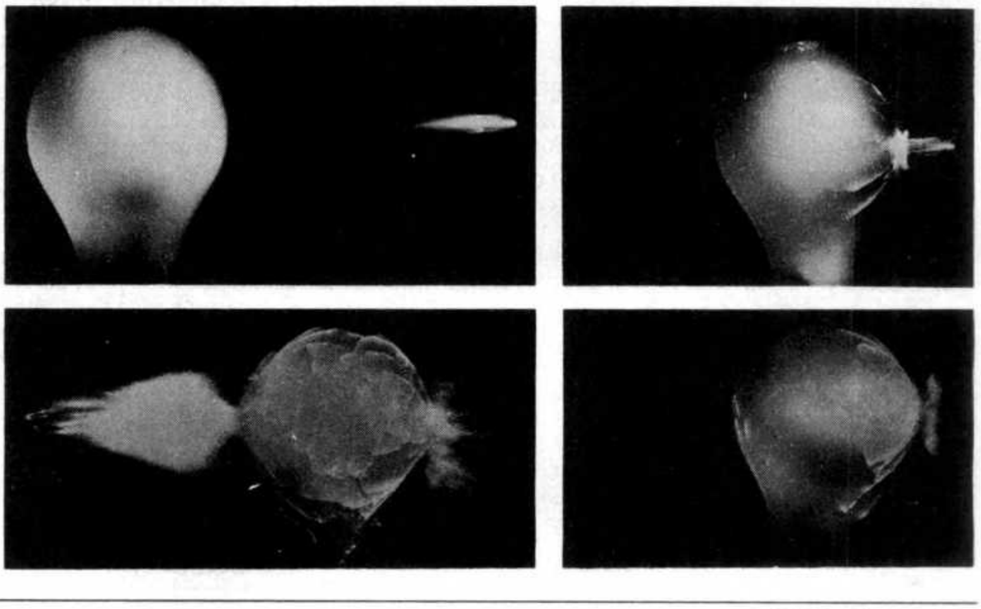
Rotating prism

It was not until 1932 that a suitable principle was developed to give an acceptable and commercial high-speed cine-camera. It was first used in the Olympic Games at Los Angeles as a photo-finish camera.

The standard cine-camera system moves film intermittently and only exposes it when it is standing still. This action severely limits the possible rates as neither the film nor the mechanism can suffer such repeated acceleration and deceleration beyond say 1,000 p.p.s. even with 16 mm film.

Allowing the film to move continuously and causing the image during exposure to move with the film produces 'image motion compensation', commonly called image compensation. This is achieved by placing a typically rectangular block of glass (prism) between the lens and the film and rotating it in synchrony with the film so that, over a useful angle, film and image move together. This method, first used successfully in 1932 in the Kirby race camera, was later taken over by Kodak, which produced the now famous Kodak high-speed camera using 16 mm film, 30 m long and taking 3,000 p.p.s. This was the standard camera in this field for almost 20 years. After 1945 it was overtaken by the Fastax camera which was lighter and produced in many models having a top speed of 8,000 full frame 16 mm p.p.s. with magazines up to 360 m. In recent years, other designs have taken over the market, usually with top speeds approaching 10,000 p.p.s. and 400-ft magazines but now with excellent speed control, synchronization and timing circuits using modern electronics.

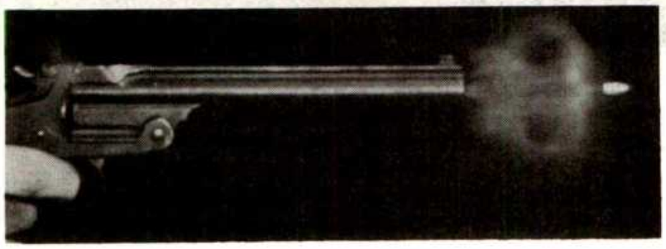
This camera class has immense value in engineering and science, making clear many problems of the motion of machines and living subjects covering perhaps 95% of all high-speed applications.



This sequence shows a 0.22 bullet piercing a lightbulb. It starts in the top left-hand frame, with the bullet approaching the bulb and (moving clockwise) ends with the bullet having passed through and the bulb beginning to collapse. Photographs by Harold E. Edgerton of MIT, Cambridge, Massachusetts.

Science Photo Library

0028–0836/81/260617—03$01.00



Subject size

Deciding the camera speed required for any study involves some prior knowledge or estimation of the subject's velocity—not its actual velocity in metres per second but its relative velocity on film or the time it would take to traverse its image across the available picture. A suitable exposure time is one-thousandth of this time and the picture rate maximum would be the reciprocal of this exposure. This maximum is reduced by the shutter factor—the ratio of exposure time to frame interval.

Thus when recording a very large subject, the camera speed can be relatively low but with a microscopic object the camera speed is very high for the same subject velocity. Thus magnification in the camera magnifies the apparent velocity.

Synchronization

With 30 m of film at 10,000 p.p.s., the camera magazine is exhausted in half a second; thus the event and the camera must be arranged to coincide to at least 0.1 s or less: this is about the limit of human reaction. Thus synchronization must be automatic, involving a knowledge of both the camera and event variables. Skill in evolving methods of synchronization is the hallmark of a top high-speed operator, although he is much assisted by the electronics built into modern cameras.

Ballistics synchro

Ballistics synchro is a specialized application of these cameras with the compensating prism removed and a stationary slit close to the film plane. The camera is placed such that the event's motion is along the length of the film and actually imitates a focal plane shutter in which the film moves instead of the shutter. The result is identical to the current photo-finish cameras in which an isolated slit of space—the finishing line—is continuously photographed, thus making the one axis along the film equal to time, that is, a continuous study is made in time of the finishing line.

This is exemplified by a projectile photographed as it passes the camera. The result when the image and camera speeds are matched is an apparently normal picture of the projectile but as is seen in photo-finish pictures in which athletes or horses appear distorted, the change of time along one axis gives an unreal but useful image, especially of subjects whose synchronization is difficult.

Streak cameras

In principle streak cameras are also the same as photo-finish cameras but differ usually in being on shorter lengths of film—for example in early designs because the film was carried on a rotating drum and its was the length of the drum's circumference. Streak cameras are used for even higher speeds than the converted prism high-speed camera whose film speed maximum was 75 m s^{-1}. Although rotating drums are not very much faster than this, we will see later techniques that are.

Picture exposure time

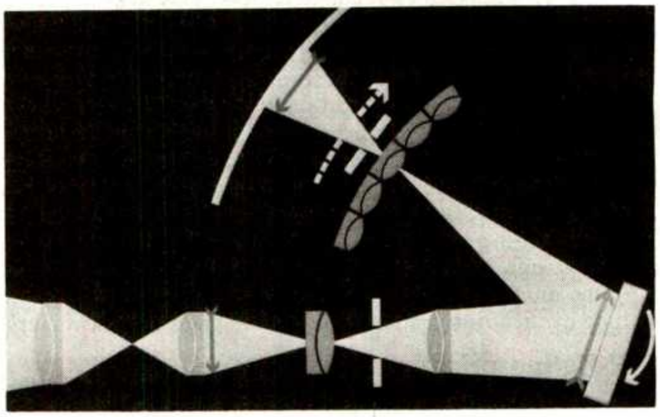
In the prism camera in its normal mode with image-compensation the maximum single picture exposure time is obviously something less than the frame interval whose minimum is 1 : 10,000 of a second or 100 μs. By inserting a slit near the film creating a stationary focal plane shutter, the individual exposure can be as short as 1 μs. In any normal single picture camera, mechanical shutters cannot readily go below 1 ms. Thus for individual pictures with very short exposures, other shutters are necessary.

There have been various 'explosive' shutters which are only capable of being used once but they are really a last resort and are only used in very specialized and unique experiments. Over the same century that high-speed cameras developed, the Kerr-cell shutter has evolved from Kerr's discovery of the electro-optic effect in 1875.

Kerr discovered that some liquids in an electric field displayed birefringence, that is, it had two different refractive indices to light vibrating in orthogonal planes, and would thus have different transit times through the same pathlength. Plane-polarized light in the correct conditions would appear to have rotated its plane of polarization while traversing the cell. (The plane polarized light will have been resolved into two beams at ±45°; one portion would have been retarded $\lambda/2$, the new combination appearing as having rotated 90°.)

We have now exchanged the motion of solid bodies—shutters, mirrors, prisms, drums—for the motion of molecules in a liquid. The final limitation occurred with solid bodies when they broke. In this case the limit is set by the molecular relaxation time typically of the order of 10^{-10} s. This cannot be fully achieved but we have reached a shutter action of 1 ns, with the flexibility of electronics to achieve any longer time. The Faraday magneto-electric effect is also usable but is electrically limited as it is necessary rapidly to change an inductive circuit as opposed to a capacitive circuit in the Kerr effect.

Both suffer from light losses as it is necessary to use polarizers giving a maximum transmission of 10% with a cutoff down by a factor of 10^{-4}. They have both been applied successfully and can give high quality images.



The rotating mirror framing camera. The film stands still while the rotating mirror moves. Image motion is compensated optically, and maximum speed (8 mm image) is about 10 million pictures per second.

Electronic flash

The first short exposure pictures (Fox-Talbot, Worthington) were by the light from a spark discharge. This was refined around 1932 by Edgerton, who put the spark in a bottle in an inert gas such as xenon and produced a brighter, more reliable light source from the same energy input. The flash duration is restricted to not less than 1 μs (in special high voltage equipment) and is in low voltage, amateur equipment about 1 ms, although the new computer flash units can be as short as 30 μs. Edgerton extended this to stroboscopes, a series of short duration flashes, with readily available units giving 2,500 flashes per second, each flash being about 1 μs, with longer, brighter flashes at lower repetition rates.

Most people who have a camera have also used an electronic flash, usually without fully appreciating its short exposure capability. To see Edgerton's pictures is really to understand what flash can do.

Sparks have been made with durations as light sources down to a few nanoseconds, the advantage being that the ionized gas which created the light can deionize and dissipate its heat sufficiently quickly whereas a discharge in gases like xenon in a bottle deionizes and dissipates more slowly.

Rotating mirror cameras

To achieve higher camera rates by mechanical means required a new means of image-compensation. Rotating mirrors to move a beam of light along a film arc had been used many times before. A good example of this was given by Blondlot (1888) who used observation through a rotating mirror to attempt to measure the time taken for the Kerr effect, his estimated minimum being 2.5×10^{-6} s, while Abraham and Lemoine (1899) measured an optical time delay of 80 cm which equals 2.7×10^{-9} s.

Miller (1939) proposed a rotating mirror framing camera system which was not really used until 1944 and even then was not totally appreciated. The basic idea is simple. You have a real image at or near the surface of a thin flat mirror which, rotating, sends the image beam to an arc of small lenses, each of which produces an image on a film arc. The light beam moves but the image detail on the film does not move because it is an image of a real image very close to the centre of rotation.

This system in a British camera (AWRE design, Barr and Stroud manufacture, JHPI distribution) can take pictures each of 8 mm diameter at rates up to 8×10^6 per s. At that rate, the mirror spins at about 5 kHz, 80% of the speed at which its own centrifugal forces would pull it apart, its outer edge moving above Mach 1. This top speed can only be marginally improved using titanium and as it was shown by Schardin (1956) that diffraction limits the practical optical beam length, these figures are as close to physical limits as it is reasonable to go.

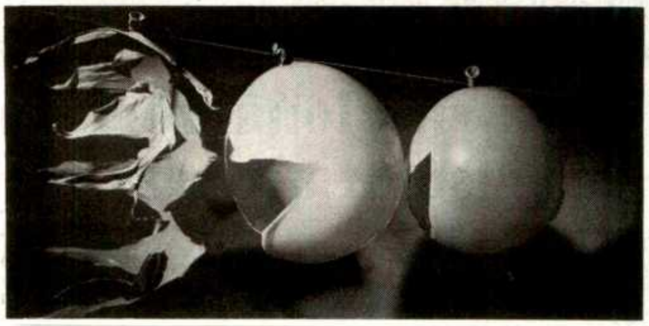
Such a rotating mirror can equally well move a slit image along the film producing the classical streak camera with a writing speed at the film of 40 mm μs^{-1} and a reasonable time resolution of 2.5 ns, which shows the excellence of nineteenth century workers. Framing and streak cameras have proved invaluable in studying explosives, properties of materials, aerodynamics, atomic fission and fusion, plasma physics and many other fields.

Having exhausted or at least reached the physical limits of solids, liquids and molecules, the only remaining hope for higher speeds was atomic particles: electrons.

Photoelectric devices

The photoelectric effect was discovered by Hertz (1887), explained by Einstein (1905) and for our purpose put to good effect by Holst and colleagues in 1934 when the first image tube was made. An image tube in its simplest form is a plane photocathode onto which an image is focused and a phosphor screen parallel and in close proximity to the photocathode with a suitable electrical potential between them such that electrons pass in direct paths one to the other and cause a visible image to appear on the screen. This is the first form to be attempted but it was not an immediate success because vacuum and clean room techniques had not been developed sufficiently.

The first good tube was produced at Mullards (1950) with a long electron path and used electromagnetic focusing. By pulsing its tube voltage (6 kV) on and off, it operated as a shutter down to 10^{-7} s and as a streak camera (a slit image at the photocathode) with electromagnetic scanning, it gave an image on film at 100 mm μs^{-1} and a possible time resolution below 1 ns. Improved technology now produces a wide variety of image tubes with proximity diodes and electrostatically scanned and focused tubes.



Another Edgerton photograph, this showing a bullet fired through a row of balloons with obvious results.

Science Photo Library

The leading camera system commercially is the Imacon (JHPI) which can be a framing camera up to 600×10^6 per s (5–15 frames) or a streak camera with resolution times of a few picoseconds. In a ballistics–synchro mode, its optical gain is enough to permit photography of a bullet in flight using only sunlight. These cameras can be made with response to IR, visual, UV or X-ray inputs. The picture output can be through a lens and camera or by contact using a fibre-optic face plate at the screen. Fibre-optic input or coupling in some image tubes is also possible.

Image intensification is possible with normal intensifiers, in cascade if desired, or even more conveniently by channel plate intensifier. This is a relatively new device (Adams and Manley, 1965) which has combined fibre-optic techniques with secondary emitters in an original manner and by a complex and clever technology with a British–French–Dutch combination through Philips. Take a fibre-optic face plate which is a mosaic of short glass rods coated by another different glass. These rods will be 40 μm diameter and 2.4 mm long, making a face plate 2.4 mm thick and 25–50 mm diameter. The glass of the rods has been selected so that it can be dissolved away leaving a face plate of tubes, the coating glass not being soluble. The inside of every tube is now coated with a secondary emitter and the end faces plated to give electrical connection. The plate is encased in an evacuated tube with a potential from face to face and a photocathode in proximity to one side and a phosphor at the other. Electrons emitted by light on the photocathode will enter the channel plate tubes and suffer numerous collisions with the secondary emitter surface, producing a shower of electrons to pass to the phosphor. Electron gains up to $\times10^5$ are possible equivalent to the same degree of image intensification and making either very faint subjects visible or giving enough signal from a very short exposure to produce a recordable image.

Coupling these devices together in various systems produces instruments which are barely credible and were certainly not imaginable a few years ago. Coupling these devices with the current advances in lasers whose very short pulses are only measurable by these newest image tubes and it is possible that we may see the dream of safe energy from fusion come true.

As almost none of these devices was reasonably imaginable at the beginning of high-speed photography, it is foolish almost to speculate. One thing is certain, we have repeatedly arrived at a physical limitation of a system. When we remember that the frequency of light is only 10^{15} Hz and lasers and image tubes are already at 10^{-12} s, there is little room for further advance except perhaps by changing to higher radiation frequencies than light. The image tube chapter is only 30 yr old, and yet time resolution has moved 4 orders (10^{-8}–10^{-12}). Could another 30 years show that even time is not divisible continuously?

Recent developments

The micro-channel plate multipliers produced under Philips with French and British cooperation have made a considerable impact in the image-tube field as they can be coupled with other tubes or incorporated as part of another tube bringing high gain intensification in an adaptable form.

In a simpler field, at camera rates above normal, are two 'convenience' systems. First the Polavision Super-8 cine-system by Polaroid–Land gives a camera capable of speeds up to 300 p.p.s. with immediate on-the-spot processing and playback 2 min after taking. It can be monochrome or colour, very portable, flexible and simpler to use—ideal for machine fault finding in automated factories.

Second, Video is now capable of leaving the real-time range and giving high speed taking with normal, still frame or low speed playback. Although not yet on the market the Spin Physics SP2000 system claims full frame rates up to 2,000 p.p.s., playback, duplication ability and *x*–*y* analysis.

Gravitational focusing and the association of distant quasars with foreground galaxies

Claude R. Canizares
Department of Physics and Center for Space Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

The surface density of quasars brighter than a given threshold is enhanced near a galaxy by the gravitational focusing of light from those quasars that are nearly co-aligned with individual stars in the galaxy halo. The focusing can intensify and make detectable quasars which would otherwise have been below the threshold. If galaxies have massive haloes of macroscopic objects then the resulting enhancement could explain the apparent association between quasars and galaxies with discrepant redshifts found in a statistical analysis. Therefore, such quasar–galaxy associations need not require the abandonment of the cosmological interpretation of the redshift.

THERE have been numerous discoveries of quasars with moderate to high redshifts located near galaxies with very small redshifts (see refs 1, 2 and references therein). Although the statistical interpretation of each of these discoveries is fraught with systematic uncertainties and is therefore controversial, the number of proximate quasar–galaxy pairs may indeed exceed that expected from the superposition of two uncorrelated populations. More global tests of entire samples give small, marginally significant indications of a positive correlation between the positions of quasars and galaxies[3,4]. The controversy surrounding all these findings is exacerbated by the implication that a positional association between galaxies and quasars necessarily implies that at least some of the quasars are at the distances of their neighbour galaxies and thus that their redshifts do not reflect the Hubble flow but require a noncosmological explanation[2].

I suggest here that the apparent surface density of quasars brighter than some threshold is in fact enhanced near a galaxy. This enhancement is due to the gravitational focusing of the light of quasars that are nearly coaligned by chance with individual stars in the galaxy halo. The focusing intensifies and makes detectable quasars whose brightness would otherwise fall below threshold. Plausible estimates of the relevant parameters imply density enhancements which are sufficient to account for the positive correlation between quasar and galaxy samples.

Gott and Gunn[5] have previously suggested that the quasar surface density enhancement due to gravitational focusing by an entire galaxy could explain a close pair of quasars; Barnothy and Barnothy[6,7] propose that all quasars are gravitationally intensified and that focusing by globular clusters enhances the quasar density near galaxies; Chang and Refsdal[8] consider the gravitational effect of a star on the quasar image formed by a galactic gravitational lens. The latter topic is also treated by Gott[9] in a paper received after my initial submission of this article. His conclusions regarding the dependence of the variability time scale on the mass of the focusing star are similar to those presented here. Finally, Turner[10] has recently considered the effects of gravitational intensification by galaxies on various measures of quasar evolution.

Gravitational focusing effects by stars on quasar surface density

The enhancement in the quasar surface density due to an assemblage of stars (point masses) is calculated in two steps. First I find the probability that a given line-of-sight through the assemblage passes close enough to a star so that its action as a gravitational lens will intensify a background light source by an amount $>\mu$. Then I fold this with the surface density of background quasars that are below the detection threshold by a factor μ and integrate over μ to derive the total enhanced density.

Einstein[11] and Refsdal[12,13] give expressions for the intensification μ due to the gravitational focusing of light by an object B of mass M_* located near the line of sight from the observer O to the source S:

$$\mu(x)=\frac{1}{2}\left[\frac{1}{x}(1+x^2)^{1/2}+x(1+x^2)^{-1/2}\right] \qquad (1)$$

where $\mu\equiv f/f_0$ and $x\equiv\beta/\alpha_0$. Here f is the flux seen by the observer, f_0 is the flux he would see in the absence of focusing, β is the angle between the true positions of B and S measured from O. Also

$$\alpha_0\simeq 4\left(\frac{GM_*}{c^2}\right)^{1/2}A_B^{-1/2}=\alpha_{11}\left(\frac{M_*}{A_B}\right)^{1/2}$$

with

$$\begin{aligned}\alpha_{11}&=8.9\times10^{-10}\quad \text{rad}\, M_\odot^{-1/2}\,\text{Mpc}^{1/2}\\ &=1.8\times10^{-4}\quad \text{s}\, M_\odot^{-1/2}\,\text{Mpc}^{1/2}\end{aligned}$$

Here A_B is the distance from 0 to B, whose redshift is assumed small. I also assume that A_B is much less than the distance from 0 to S and that the splitting of the image is unresolved.

From equation (1) μ is >1.06 for $x<1.0$ or $\beta<\alpha_0$. Thus the light of a quasar will be noticeably intensified if it falls within a solid angle $\sim\pi\alpha_0^2=1.0\times10^{-7}(M_*/M_\odot)\ (A_B/1\ \text{Mpc})^{-1}\ \text{s}^2$ of a star of mass M_* at distance A_B.

The effect of gravitational focusing on the apparent surface density of quasars near galaxies is calculated as follows. The differential surface density of stars with mass M_* in a galaxy is given by the expression $(\mathrm{d}N_*/\mathrm{d}\Omega)\psi(M^*)\,\mathrm{d}M$ where $\mathrm{d}N_*/\mathrm{d}\Omega$ is the surface density of all stars at some point in the galaxy and $\psi(M_*)$ is the mass distribution function, which I assume is the same throughout the galaxy. Then the probability density that a line of sight from Earth through the galaxy passes within x to $x+\mathrm{d}x$ of a star with mass between M_* and $M_*+\mathrm{d}M_*$ is

$$p(x,M_*)\,\mathrm{d}x\,\mathrm{d}M_*=2\pi\alpha_{11}^2A_B^{-1}\left(\frac{\mathrm{d}N_*}{\mathrm{d}\Omega}\right)M_*\psi(M_*)\,\mathrm{d}M_*x\,\mathrm{d}x$$

Integrating over all M_* gives

$$p(x)\,\mathrm{d}x=2\pi\alpha_{11}^2A_B^{-1}\frac{\mathrm{d}N_*}{\mathrm{d}\Omega}\langle M_*\rangle x\,\mathrm{d}x \qquad (2)$$

where $\langle M_*\rangle$ is the mean stellar mass in the galaxy. Equations (1) and (2) can be used to transform $p(x)$ to $p(\mu)$, which can then be integrated to derive the probability $P(\mu)$ that a given line of sight will result in an intensification $>\mu$:

0028–0836/81/260620—05$01.00

$$P(\mu)=\int_{\mu}^{\infty} p(\mu)\,\mathrm{d}\mu=\frac{\pi}{2}\frac{\alpha_{11}^{2}}{A_{\mathrm{B}}}\Sigma_{M}[\mu(\mu^{2}-1)^{-1/2}-1]$$

where $\Sigma_M \equiv \langle M_* \rangle \, \mathrm{d}N_*/\mathrm{d}\Omega$ is the surface density of mass (per solid angle) in the galaxy measured along the line of sight.

$\Phi_\theta(f_t)$ is the angular surface density of quasars with apparent flux greater than a threshold f_t at some angular distance θ from the centre of a galaxy that has an outer radius R and is at a distance A_{B}. The enhanced density is related to the surface density of quasars brighter than f_t in the absence of focusing, $\Phi_0(f_t)$, by the expression

$$\Phi_\theta(f_t)=\Phi_0(f_t)+\int_0^{f_t}\Phi_0'(f)P\left(\frac{f_t}{f}\right)\mathrm{d}f \qquad (3)$$

for $\theta < R/A_{\mathrm{B}}$. Here $\Phi_0'(f)$ is the derivative of $\Phi_0(f)$. The second term gives the enhancement due to the gravitational intensification of quasars which would otherwise fall below the detection threshold.

I take a quasar surface density of the form $\Phi_0(f)=\phi_0(f/f')^{-\eta}$ for $f > f_{\min}$ where η, ϕ_0 and f' are constants. Below I will piece together a surface density distribution from several segments of this form. The integration of equation (3) then yields

$$\Phi_\theta(f_t)=\Phi_0(f_t)[1+\zeta(\theta)] \qquad (4)$$

where

$$\zeta(\theta)=F(\eta;\mu_{\max})(\pi/2)\alpha_{11}^{2}A_{\mathrm{B}}^{-1}\Sigma_M(\theta) \qquad (5)$$

Here $F(\eta;\mu_{\max})$ is an expression which generally depends on both η and $\mu_{\max}=f_t/f_{\min}$, the intensification required to bring the faintest quasars above threshold. However, the $\mu_{\max}$ dependence is important only when $\eta > 2$, because then the rise in the surface density of faint quasars compensates the fall in the probability of large intensification, so the faintest quasars dominate. Where $F(1;\mu_{\max})\simeq 1$ and $F(1.5,\mu_{\max})\simeq 2.3$, one finds $F(2;\mu_{\max})=1/2+\ln(2\mu_{\max})$, and $F(2.5;\mu_{\max})\simeq(5/2)(\mu_{\max}^{1/2}-0.5)$ for $\mu_{\max}\geqslant 1.5$.

Equation (5) can be recast in simple terms if the mass distribution of the galactic halo approximates a self-gravitating isothermal sphere out to some radius R. Then away from the core but inside R one has $\Sigma_M(\theta)\simeq A_{\mathrm{B}}\sigma_{\mathrm{v}}^2/2G\theta$, where σ_{v} is the line-of-sight velocity dispersion in the halo[9]. This gives

$$\zeta(\theta)=\frac{\pi F\alpha_{11}^{2}\sigma_{\mathrm{v}}^{2}}{4G\theta}$$

$$=4\pi F\left(\frac{\sigma_{\mathrm{v}}}{c}\right)^{2}\frac{1}{\theta}, \quad \text{for } \theta < R/A \qquad (6)$$

Finally, with equation (6), equation (4) can be integrated over a solid angle $\Omega=\pi\theta_{\max}^2$ centred on the galaxy to find $\langle n\rangle_\Omega$, the expected number of quasars that are above the threshold within the solid angle, in terms of $\langle n\rangle_0$, the expected number in the absence of gravitational focusing.

$$\langle n\rangle_\Omega \simeq \langle n\rangle_0\left[1+\frac{1}{\Omega}\int\zeta(\theta)\,\mathrm{d}\Omega\right]$$

$$\simeq\langle n\rangle_0[1+2\zeta(\theta_{\max})], \quad \text{for } (\theta_{\max} < R/A_{\mathrm{B}}) \qquad (7)$$

with $\zeta(\theta)$ given by equation (6).

To estimate parameters I first consider $F(\eta;\mu_{\max})$. The value of η is found by comparing measurements of quasar surface densities at different limiting magnitudes. These measurements are generally uncertain by up to a factor of 10 (refs 15, 16) but there is some indication that η varies from $\eta \sim 2.5$ at limiting magnitudes ~16–18 mag (ref. 17), through $\eta \sim 2.0$ at 19 mag (ref. 18) to $\eta \sim 1.5$ at ~21–22 mag (refs 15, 16, 18, 19 and M. Schmidt, personal communication). Taking a piecewise continuous quasar surface density with these values of η and breaks at 18 and 21 mag gives the dependence of F on limiting magnitude shown in Fig. 1.

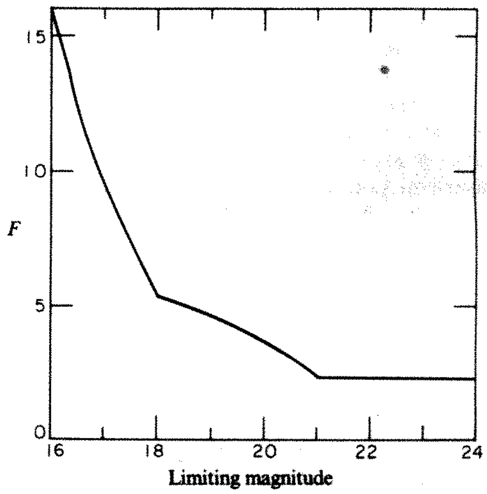


Fig. 1 Approximate value of the factor F, which appears in expressions for the enhanced quasar surface density, compared with the limiting quasar magnitude. Here F has been computed for a piecewise continuous unenhanced quasar surface density of the power-law form given in the text with exponent $\eta = 2.5$ for limiting magnitudes $m < 18$, $\eta = 2.0$ for $18 < m < 21$, and $\eta = 1.5$ for $m > 21$.

The line-of-sight velocity dispersion, σ_{v}, in the outer regions of a galaxy halo can be estimated from measured velocity dispersions and from rotation curves using assumed distributions of the galactic mass[20–22]. In the nuclear regions of elliptical galaxies one observes dispersions of 200–400 km s^{-1} (ref. 23); the highest values are for the most luminous galaxies. In the absence of additional information one assumes that the halo and nuclear velocity dispersions are similar, although Dressler[24] found that for a cD galaxy the dispersion rises to ~500 km s^{-1} in the envelope. For the spiral galaxies, measurements of bulge velocity dispersions and of rotation curves yield estimates of $\sigma_{\mathrm{v}} < 250$ km s^{-1} for the haloes[20,22]; again σ_{v} increases with galaxy luminosity.

Scaling equation (6) to representative values of the parameters gives

$$\zeta(\theta)=0.22\left(\frac{F}{5}\right)\left(\frac{\sigma_{\mathrm{v}}}{300\ \mathrm{km\ s^{-1}}}\right)^{2}\left(\frac{\theta}{\mathrm{min}}\right)^{-1} \qquad (8)$$

Thus if galactic haloes are made of macroscopic objects, the surface density of quasars seen through the haloes can be significantly enhanced. For bright quasars and massive galaxies one can have values of ζ four or five times larger than the representative value in equation (8), giving from equation (7) a doubling of the expected number of quasars within 2 arc min of the galaxy. Of course, our ignorance about the true quantity and distribution of mass in galactic haloes and about the surface densities of faint quasars means that the numerical values of ζ should be treated as approximate estimates of the true enhancement factor.

The above analysis neglects the gravitational focusing due to the distributed, total mass of the galaxy. One can show from the equations of refs 5, 9–14 that the enhancement in quasar surface density for the galaxy, ζ_{g}, depends on the 'critical angle' $\theta_{\mathrm{c}} \simeq 8\sigma_{\mathrm{v}}^2/c^2$. For $\sigma_{\mathrm{v}} = 300$ km s^{-1}, $\theta_{\mathrm{c}} = 1.7''$. At angles θ such that $R/A_{\mathrm{B}} > \theta \gg \theta_{\mathrm{c}}$ one has $\zeta_{\mathrm{g}}(\theta) \simeq (\eta-1)\theta_{\mathrm{c}}/\theta$, where η is the exponent in the quasar surface density distribution at the limiting magnitude of interest (because the intensification by the galaxy as a whole is small for $\theta \gg \theta_{\mathrm{c}}$, the enhancement is due to quasars only slightly fainter than the threshold). If $\eta = 1$, the rise in the surface density of faint quasars just compensates the loss in solid angle caused by the magnification of the galaxian gravitational lens, and $\zeta_{\mathrm{g}} = 0$. The ratio of $\zeta_{\mathrm{g}}(\theta)$ to $\zeta(\theta)$ for stellar focusing from equation (6) is

$$\frac{\zeta_{\mathrm{g}}}{\zeta}=\frac{2(\eta-1)}{\pi F}$$

This is never greater than ~0.15 for the parameters specified above, so ζ_{g} can be safely neglected in the present work. When

$\theta \sim \theta_c$, the galaxian focusing can dominate; this is the relevant regime for discussion of the double quasar[14] and is considered by Chang and Refsdal[8] and Gott[9]. Note that a globular cluster has $\sigma_v < 10\ \mathrm{km\ s^{-1}}$ (ref. 25) and $\theta_c < 0.002''$ so it has a negligible effect on quasar surface densities, contrary to the claim of ref. 7.

It is also necessary to consider the importance of the finite size of the quasar emission region (the finite size of the intervening star has no effect for $M_* \gg (10^6 \rho^{-2} A_B^{-3})$g, where ρ is its density in g cm^{-3} and A_B is in Mpc). The size scale at the quasar is given by $l_0 \sim A_s \alpha_0$. Taking A_s, the 'angular diameter distance' to S, from the expressions of Young *et al.*[14] gives

$$l_0 \sim 0.4 \left(\frac{H_0}{100}\right)\left(\frac{M_*}{M_\odot}\right)^{1/2}\left(\frac{A_B}{10\ \mathrm{Mpc}}\right)^{-1/2}\ \mathrm{pc}$$

for quasar redshifts >0.5. Here H_0 is Hubble's constant in km s^{-1} Mpc^{-1}.

The dimensions of the continuum emitting regions of quasars are thought to be $\ll l_0$ (ref. 26), so equations (5) and (6) are directly applicable to them. The regions of broad-line emission probably have dimensions $\sim l_0$ unless $M_* \ll 1\ M_\odot$. One can show that an extended source of uniform surface brightness and dimension l will undergo an intensification $\mu_e \sim [1 + (l_0/l)^2]^{1/2}$ if it is co-aligned with the star. As it is possible to have $\mu > \mu_e$, the broad emission lines of quasars with large continuum region intensifications could appear anomalously weak. The dimensions of the narrow line emitting regions of quasars are almost certainly $> l_0$ (ref. 26), and so these lines would undergo negligible intensification as would radio or IR emission from extended regions.

The mass M_* of the individual objects serving as gravitational lenses enters into the expression for l_0 and for the variability time scale discussed below but does not appear explicitly in the quasar surface density enhancement of equation (5). The latter equation depends only on the mass surface density in the intervening galaxy. (In fact, one can use equation (5) to recover the order of magnitude of Press and Gunn's[27] result that the enhancement in quasar surface density for a universe with a uniform distribution of point masses is approximately equal to Ω_0, the ratio of the density of the point masses to the cosmologically critical density; see also ref. 9). The validity of the single scattering approximation on which equation (5) is based is also independent of M_* (the probability that a line of sight falls within $\pi\alpha_0^2$ of a star in a galaxy halo is $8\pi\sigma_v^2/c^2\theta = 0.09$ for $\sigma_v = 300\ \mathrm{km\ s^{-1}}$ and $\theta = 1'$). However, equation (1) is based on geometrical optics[11,28] which requires that the individual objects have dimensions that are large compared with the wavelength of the focused radiation. If galactic haloes are composed largely of massive neutrinos, then they would not contribute to equation (5) and the value of ζ would be substantially reduced.

Comparison with observations

I now compare the estimates of the surface density enhancement of quasars near galaxies with the observational evidence of Seldner and Peebles[3] and Burbidge[1]. This is done by deducing from the observational data those values of ζ which would account for the apparent quasar–galaxy correlations.

Seldner and Peebles[3] find a statistically significant enhancement in the galaxy–quasar two-point angular correlation function. The enhancement is equivalent to that which would occur if there were 570 ± 150 extra galaxies added to the ~5,300 galaxies expected by chance within 17.5′ of all the quasars in the sample. Their result has the virtue of being completely independent of the assumed value of the mean quasar surface density, although it is certainly affected by selection effects as they discuss.

Because of the galaxy–galaxy two-point correlations[29], a real association between a quasar and a single galaxy will give rise to an apparent correlation between the quasar and other galaxies as well. Within some distance δ of the quasar (and its associated galaxy) one would find an excess number of galaxies given by

$$N_e \simeq 1 + n_g \int_0^{\delta} (w_{gg}(\delta) - 1)\, d\Omega(\delta)$$

where n_g is the mean galaxy density[30] and w_{gg} is the galaxy–galaxy two-point correlation function[29]. Evaluating this for $\delta = 17.5'$ gives $N_e = 75$ so that only ~8 of the 382 quasars in the Seldner and Peebles sample need have a real association with a galaxy to account for the excess quasar–galaxy correlation. This is consistent with Seldner and Peebles' finding that the enhanced quasar–galaxy correlation disappears when the 10 largest contributors to the enhancement are removed. In fact, 3 of those 10 are on Burbidge's list of quasar–galaxy pairs, which further suggests that the excess correlation could be due to a small number of 'real' associations.

It is possible to make an estimate of $\bar{\zeta}_{sp}$, the implied value of the average quasar density enhancement factor for the galaxies in Seldner and Peebles' sample that is needed to account for the ~8 true quasar–galaxy associations. The total number of observed quasars N_Q is given by

$$N_Q \cong N(1 + 2\bar{\zeta}_{sp}\Omega_G/\Omega_T) \qquad (9)$$

where N is the number in the absence of focusing and Ω_G/Ω_T is the fractional solid angle subtended by galaxies. The typical galaxy in the Shane–Wirtanen sample used by Seldner and Peebles has redshift ~0.08 (ref. 3) so $A_B \sim 240$ Mpc for $H_0 = 100\ \mathrm{km\ s^{-1}\ Mpc^{-1}}$. The galaxy luminosity function in a sample limited by apparent magnitude is peaked near $L^* = 2.5 \times 10^{10}\ L_\odot$ (ref. 31) and so has $\sigma_v \sim 300\ \mathrm{km\ s^{-1}}$ (ref. 23). The outer radius R of the halo of such a galaxy is probably ~3 times its Holmberg radius of 40 kpc (refs 21, 32). Thus a typical galaxy subtends 2.4×10^{-3} deg^2, and the mean galaxy density is ~50 deg^{-2} (ref. 30), so $\Omega_G/\Omega_T \sim 0.1$. Also $N_Q = 380$, $N \leq N_Q$, and I require $2N\bar{\zeta}_{sp}\Omega_G/\Omega_T \sim 8$. Inverting this last expression gives $\bar{\zeta}_{sp} \sim 0.1$. This compares remarkably well to equation (8), which with nominal values for F, and σ_v, and with $\theta = R/A_B \sim 1.7'$ gives $\zeta = 0.13$.

Burbidge[1] has made a statistical analysis of specific galaxy–quasar associations found by Arp and others. For separations ≤3′ there is an excess of observed over predicted quasars nearly all of which is due to the excess between 1′ and 2′. For quasar–galaxy separations of 3′–10′ there is a deficit of observed quasars; presumably selection effects have caused many to be overlooked. The predicted number is computed using a quasar surface density deduced from several studies of quasars in the field and using an estimate of the number of regions sampled. Both these numbers are very uncertain (see ref. 15), and the whole analysis has been questioned[33].

If one takes from Table 2 of Burbidge the observed (N_0) and his estimate of the expected ($\langle n \rangle_0$) number of quasars within some angular separation (θ_{max}) then one can set $\langle n \rangle_\Omega = N_0$ in equation (7) and compute the mean enhancement factor $\bar{\zeta}_B(\theta_{max})$ required if gravitational focusing is to account for the entire excess, $\bar{\zeta}_B = (N_0/\langle n \rangle_0 - 1)/2$. Table 1 lists the data both for the full quasar–galaxy sample and for the subset discovered by Arp. Note that the table entries are cumulative both across and down and so are not statistically independent.

Burbidge's computation of $\langle n \rangle_0$ is probably more applicable to the Arp subset[1], so it is the values of $\bar{\zeta}_B$ from the upper portion

Table 1 Mean quasar density enhancement for Burbidge's compilation

	$\theta_{max} = 1'$			$\theta_{max} = 2'$			$\theta_{max} = 3'$		
Mag	N_0	$\langle n \rangle_0$	$\bar{\zeta}_B$	N_0	$\langle n \rangle_0$	$\bar{\zeta}_B$	N_0	$\langle n \rangle_0$	$\bar{\zeta}_B$
Arp sample									
<18	0	0.17	—	4	0.7	2.3	5	1.6	1.1
<19	1	0.52	0.46	10	2.1	1.9	12	4.6	0.8
<20	3	1.73	0.37	15	6.9	0.6	19	15.5	0.1
Total sample									
<18	5	0.17	14.0	11	0.7	7.3	12	1.6	3.3
<19	9	0.52	8.2	21	2.1	4.5	25	4.6	2.2
<20	12	1.73	3.0	28	6.9	1.5	34	15.5	0.6

N_0 is the observed number of quasars and $\langle n \rangle_0$ the predicted number (for Arp sample—see text) from Burbidge[1].

of Table 1 that might be compared with equation (8). These $\bar{\zeta}_B$s exceed our nominal estimate of ζ by factors of three or more. Unfortunately, these $\bar{\zeta}_B$s have little or no significance because of their large statistical errors (most of the contribution to the $\bar{\zeta}_B$s comes from ~8 too many quasars brighter than 19 mag with $1' < \theta_{max} < 2'$) combined with their sensitivity to unknown selection effects. Further data would be required to establish a real discrepancy.

The very large values of $\bar{\zeta}_B$ computed for the total quasar–galaxy sample cannot be taken seriously because the predicted $\langle n\rangle_0$s are surely underestimated for this case. However, it is interesting that the relative values of $\bar{\zeta}_B$ in both portions of Table 1 are qualitatively in agreement with equation (8): for a given limiting magnitude, $\bar{\zeta}_B$ does in fact go roughly as $\sim\theta_{max}^{-1}$; for a given θ_{max}, $\bar{\zeta}_B$ does fall as the limiting magnitude increases from 18 to 19 in approximate accord with the fall in F of Fig. 1. It is unlikely that the quasar sample is complete to 20 mag, so it is not too surprising that the same scaling fails when one compares $\bar{\zeta}_B$s for magnitudes 19 and 20.

Burbidge also notes that the quasar–galaxy angular separation is inversely correlated with the galaxy redshift. This is consistent with the hypothesis of enhancement by gravitational focusing because for fixed R, the angular size of the region of enhanced quasar surface density would fall with increasing galaxy redshift.

Burbidge[1] and Arp[2] list several special cases of multiple quasar–galaxy associations. The calculation of the statistical significance of these cases is very difficult because of the unknown sample size and selection criteria. Furthermore, their significance is very sensitive to the assumed quasar surface density because one always has $\langle n\rangle_0 \ll 1$ for which the Poisson probability of detecting i quasars goes as $\langle n\rangle_0^i$. Of course, this also means that for quasars observed through the halo of a galaxy, the probability goes as $\langle n\rangle_\Omega^i \sim \langle n\rangle_0^i(1+2\zeta(\theta))^i$ (from equation (7)), so that even modest values of ζ could have a significant effect on the inferred significance.

There are several special cases for which a density enhancement is not a sufficient explanation. These include quasar pairs with similar redshifts, which imply clustering or super-clustering (see refs 15, 34, 35) and quasars with supposedly unlikely orientations or alignments[1,2] (but see ref. 36).

Conclusions

I have shown that the quasar surface density can be significantly enhanced by the chance alignment of faint quasars with individual stars in the haloes of foreground galaxies, stars that act as gravitational lenses to intensify the quasar light. The estimated magnitude of the effect, which has a large uncertainty due to our poor knowledge of the relevant parameters, is sufficient to explain fully the association of distant quasars with foreground galaxies found in the statistical analysis of Seldner and Peebles[3]. With nominal parameters the effect falls short of Burbidge's[1] estimates of the enhanced quasar density derived from Arp's quasar–galaxy associations. However, these estimated enhancements have little or no significance when statistical and systematic uncertainties are considered. Burbidge's estimates of quasar density enhancements near galaxies derived from his full catalogue do have approximately the dependences on angle and limiting magnitude that one would expect if the enhancements were due to gravitational focusing.

I conclude that the identification of this physically plausible mechanism for obtaining galaxy–quasar associations of the appropriate magnitude undermines the claims that such associations must be the symptoms of non-cosmological redshifts for some quasars.

There are several observations which could test the significance of gravitational focusing by stars on the quasar surface density. Clearly, an increased quasar sample with well defined statistical properties is extremely important. The growing sample of X-ray selected quasars from the Einstein Observatory may constitute such a set. Tests of this sample by the techniques of Seldner and Peebles[3] would have to be consistent with the numerical estimates of ζ computed with what one hopes will be continually improved galaxy parameters. Such an exercise could establish the existence of massive galaxy haloes composed of macroscopic objects, an existence already suggested by the agreement with the Seldner and Peebles data.

In addition to statistical tests, one can search for pecularities in individual objects indicative of gravitational focusing. First, consider long-term variability of the quasar caused by the relative motion of the Earth and the star acting as the gravitational lens. Stellar velocities are ~200–300 km s^{-1}, and the galaxy may have some transverse bulk velocity relative to Earth. The impact parameter b of an undeflected quasar light ray hitting Earth is between $b = b_0 \sim A_B\alpha_0 = 4(GM_*A_B/c^2)^{1/2}$, and $b \geqslant 0.05b_0$. If $M_* = 1\,M_\odot$ and $A_B \sim 10$ Mpc then $b_0 \sim 9\times10^{15}$ cm, and one would expect the quasar and star to become misaligned in $\sim 30(b/b_0)(v/100\text{ km s}^{-1})^{-1}$ yr, where v is the component of the stellar velocity perpendicular to the line-of-sight and in the Earth–star–quasar plane. The misalignment will cause the previously magnified quasar to fade once and for all typically by several magnitudes. If none of the Burbidge objects showed this behaviour over a ~20-yr period my conclusion would be untenable (unless $M_* \gg 1\,M_\odot$). Conversely, if it can be shown that intensification by gravitational focusing does occur for some of Burbidge's objects, that they persist for >1 yr already implies that M_* is not $\ll 10^{-3}\,M_\odot$; with a significant sample of objects it might someday be possible to estimate a mean value for M_* (this suggestion is also made by Gott[9]).

Another possible test concerns the short-time scale variability of the quasar. There will be a time delay of $\Delta t \sim \alpha_0 b_0/c = 16GM_*/c^3$ between light signals arriving at Earth from a point in the quasar along the two possible trajectories that pass the star[13]. Any variations in the quasar signal, whether inherent or due to scintillations along the path from the quasar to the star, would cause an enhancement in the autocorrelation function of the signal at a lag time Δt. Unfortunately, for $M_* = 1\,M_\odot$, $\Delta t \sim 80$ µs, so this is an exceedingly difficult measurement. Nevertheless, a positive detection would validate the hypothesis and measure M_* directly.

Finally, if the dimensions of the broad line emission region are $> l_0$, then the lines would generally be intensified less than the continuum. In this case, intensified quasars would have anomalously weak lines, but the lines would fade less slowly than the continuum as the quasar and star became misaligned. Any quasar showing narrow line emission is not being intensified unless $M_* \gg 1\,M_\odot$.

I thank George Clark, Irwin Shapiro and Julian Krolik for helpful comments. Julian Krolik first suggested that intensified quasars might have weak emission lines. I received helpful comments on the initial manuscript from C. Arp, J. Bahcall, J. R. Gott and M. Schmidt. My attention was drawn to the question of quasar-galaxy associations during the 1980 Astrophysics Workshop at the Aspen Center for Physics; I thank the Center for its hospitality and NASA for its support of the workshop. I thank Brenda Parsons and Maggie Carracino for preparing the manuscript. This work was supported in part by NASA contract NAS-8-30752. I am an Alfred P. Sloan Foundation Fellow.

Received 10 March; accepted 13 May 1981.

1. Burbidge, G. *Nature* **282**, 451–455 (1979).
2. Arp, H. *Ann. N.Y. Acad. Sci.* **336**, 94–112 (1980).
3. Seldner, M. & Peebles, P. J. E. *Astrophys. J.* **227**, 30–36 (1979).
4. Roberts, D. & O'Dell, S. *Astr. Astrophys.* **76**, 254–256 (1979).
5. Gott, J. R. & Gunn, J. E. *Astrophys. J. Lett.* **190**, L105–L108 (1974).
6. Barnothy, J. M. & Barnothy, M. F. *Astrophys. J.* **174**, 477–482 (1972).
7. Barnothy, J. *Bul. Am. astr. Soc.* **6**, 212 (1974).
8. Chang, K. & Refsdal, S. *Nature* **282**, 561–564 (1979).
9. Gott, J. R. *Astrophys. J.* **243**, 140–146 (1981).
10. Turner, E. L. *Astrophys. J. Lett.* **242**, L135–140 (1980).
11. Einstein, A. *Science* **84**, 506–7 (1936).
12. Refsdal, S. *Mon. Not. R. astr. Soc.* **128**, 295–306 (1964).
13. Refsdal, S. *Mon. Not. R. astr. Soc.* **132**, 101–111 (1966).
14. Young, P., Gunn, J. E., Kristian, J., Oke, J. & Westphal, J. *Astrophys. J.* **241**, 507–520 (1980).
15. Wills, D. *Phys. Scr.* **17**, 333–337 (1978).
16. Gilmore, G. *Mon. Not. R. astr. Soc.* **195**, 183–196 (1981).
17. Green, R. F. & Schmidt, M. *Astrophys. J. Lett.* **220**, L1–L4 (1978).
18. Bohuski, T. & Weedman, D. *Astrophys. J.* **231**, 653–658 (1979).

19. Bahcall, J. & Soniera, R. M. *Astrophys. J. Lett.* **238,** L17–L20 (1980).
20. Gott, J. R. *An. Rev. Astr. Astrophys.* **15,** 235–266 (1977).
21. Faber, S. M. & Gallagher, J. S. *A. Rev. Astr. Astrophys.* **17,** 135–188 (1979).
22. Whitmore, B., Kirshner, R. & Schechter, P. *Astrophys. J.* **234,** 68–75 (1979).
23. Faber, S. M. & Jackson, R. E. *Astrophys. J.* **204,** 668–683 (1976).
24. Dressler, A. *Astrophys. J.* **231,** 659–670 (1979).
25. Peterson, C. J. & King, I. R. *Astr. J.* **80,** 427–436 (1975).
26. Davidson, K. & Netzer, H. *Rev. Mod. Phys.* **51,** 715–766 (1979).
27. Press, W. H. & Gunn, J. E. *Astrophys. J.* **185,** 397–412 (1973).
28. Weinberg, S. *Gravitation and Cosmology,* 191 (Wiley, New York, 1972).
29. Groth, E. J. & Peebles, P. J. E. *Astrophys. J.* **217,** 385–405 (1977).
30. Seldner, M., Siebers, B., Groth, E. & Peebles, P. J. E. *Astr. J.* **82,** 249–256 (1977).
31. Schechter, P. *Astrophys. J.* **203,** 297–306 (1976).
32. Holmberg, E. *Galaxies and the Universe* (eds Sandage, A. & Kristian, J.) 154–155 (University of Chicago Press, 1975).
33. Weedman, D. W. *Astrophys. J.* **237,** 326–330 (1980).
34. Burbidge, E. M., Junkkarinen, V. T., Koski, A. T., Smith, H. E. & Hoag, A. A. *Astrophys. J. Lett.* **242,** L55–L59 (1980).
35. Oort, J. H., Arp, H. & de Ruiter, H. Preprint No. 129 (ESO, 1980).
36. Edmunds, M. G. & George, G. H. *Nature* **290,** 481–483 (1981).

Sunspot periods in the late Precambrian glacial climate and solar–planetary relations

G. E. Williams

Exploration Department, The Broken Hill Proprietary Company Limited, PO Box 559, Camberwell, Victoria 3124, Australia

Cycles conspicuous in late Precambrian ($\sim$680 Myr) glacial varves in Australia reveal strong climatic periods of about 11, 22, 145 and 290 yr and a weaker period near 90 yr. The 11-, 22- and 90-yr cycles equate with sunspot periods, and the longer cycles with solar and climatic periods as indicated by tree-ring studies. The varve cycles provide clear evidence of 11- and 22-yr climatic periods in the geological record and may carry important implications for solar physics and planetary science.

THE relation between sunspot cycles and patterns of terrestrial weather and climate remains controversial[1–3]. The 11- and/or 22-yr solar periods have long been perceived in numerous meteorological and climatic variables[2,4,5], and recent studies have identified a 22-yr period in tree rings[6,7]. Others, however, suggested[8–11] that little or no convincing evidence exists for statistically significant correlations of either the 11- or 22-yr sunspot cycle with variations of weather and climate.

With regard to the geological record, periods of $\sim$11- and $\sim$22-yr and more, perceived in certain 'varved' (annually layered) sequences principally of Phanerozoic age, have been ascribed a solar origin[12–16]. With rare exception[16], however, such cyclicity is not conspicuous but is revealed only through the statistical processing of varve thickness. In the absence of a convincing explanation of solar-cycle influence on weather or of clear empirical evidence for such cyclicity in the climatological or geological record, strong doubts remain about the reputed terrestrial effects of sunspot variability.

Late Precambrian varve cycles described here provide the best evidence yet of 11- and 22-yr climatic periods in the remote past. The presence of such periods in rocks hundreds of millions of years old may have important implications for solar physics and planetary science.

Late Precambrian cyclic lamination

A cyclic lamination is conspicuous in the late Precambrian Elatina Formation at Pichi Richi Pass [32°25′ S, 137°59′ E] and at other localities up to $\sim$30 km distant in the Flinders Ranges, South Australia. There the formation comprises mainly red and brown siltstones and fine sandstones with a thickness of $\sim$30–60 m (refs 17–19). The Elatina Formation is a distal facies of western provenance deposited on a western 'shelf' of the Adelaide Geosyncline during the Marinoan glaciation[17,19]. Correlative proximal outwash sandstones occur on the shelf $\sim$100–150 km north-west of Pichi Richi Pass, indicating the former presence of an ice sheet on the low-relief cratonic region west of the shelf. Marinoan glaciation occurred $\sim$680 Myr ago[20] in low palaeolatitudes[21].

The cyclic lamination (Fig. 1) is characterized by the regular repetition of dark red brown, clayey bands spaced 2–16 mm apart that bound groups of paler, clastic laminae. The dark bands typically are 0.1–0.5 mm thick and the clastic laminae 0.1–2.5 mm thick. The lamination is essentially planar and laterally persistent over several hundred metres within the limits of outcrop.

The clastic laminae typically comprise a lower, paler, sandy layer and an upper, silty layer. The contact between such layers is usually gradational, whereas contacts between adjacent laminae are relatively sharp. Microscopically, the clastic laminae consist mainly of grains of quartz (dominant) and feldspar in a clayey matrix. Grading results from an upward reduction in grain size from 20–100 μm (silt to very fine sand) at the base to 20–50 μm (silt) at the top of each lamina and an accompanying upwards increase in the proportion of matrix. A few laminae (usually the thickest lamina in a particular group) contain one or two thin, impersistent layers of silt or very fine sand. Such layers are regarded as internal features, rather than discrete clastic laminae, because of their impersistence and thinness compared with adjacent laminae. The dark bands comprise a fine-grained mixture of clayey material and iron oxide. Some bands that appear paler are seen under magnification to consist of several very thin ($\leq$0.2 mm) clayey layers interleaved with silty laminae. Such a composite dark band occurs between cycles 9 and 10 in Fig. 1*a*.

The inspection of outcrops and sawcuts suggested that a uniform number of clastic laminae (about 10) occur in each group or 'cycle' between successive dark bands. Thin sections (Fig. 1*a*), however, reveal additional, very thin ($\leq$0.2 mm) laminae of silt at the bottom and/or top of most cycles. Sixteen successive relatively thick cycles so examined contain 9–14 discrete clastic laminae, and average 11.3 ± 1.6 (Table 1). Two additional sequences each of 10 successive cycles studied in thin section average 11.2 ± 1.7 (range 9–14) and 11.0 ± 1.0 (range 9–12) clastic laminae per cycle. Although counts for several cycles have uncertainties of ±1 because of the indistinctness of some laminae, the results overall are consistent and demonstrate the regularity of the cyclicity. The mean for the 36 cycles so examined is 11.2 laminae, and this value is used here to estimate laminae numbers for long sequences. Thin cycles from other stratigraphic levels of the Elatina Formation (Fig. 1*b*, *c*) seem to contain about the same number of clastic laminae as the thick cycles, but the thinness of laminae hinders accurate counting.

The thickness of clastic laminae varies systematically, attaining a maximum near the centre of each cycle (Figs 1*a* and 2). Most cycles show some asymmetry, as the thickest lamina usually occurs below or above the lamina of median number. Of the cycles listed in Table 1, nine show positive skewness (for

0028-0836/81/260624—05$01.00

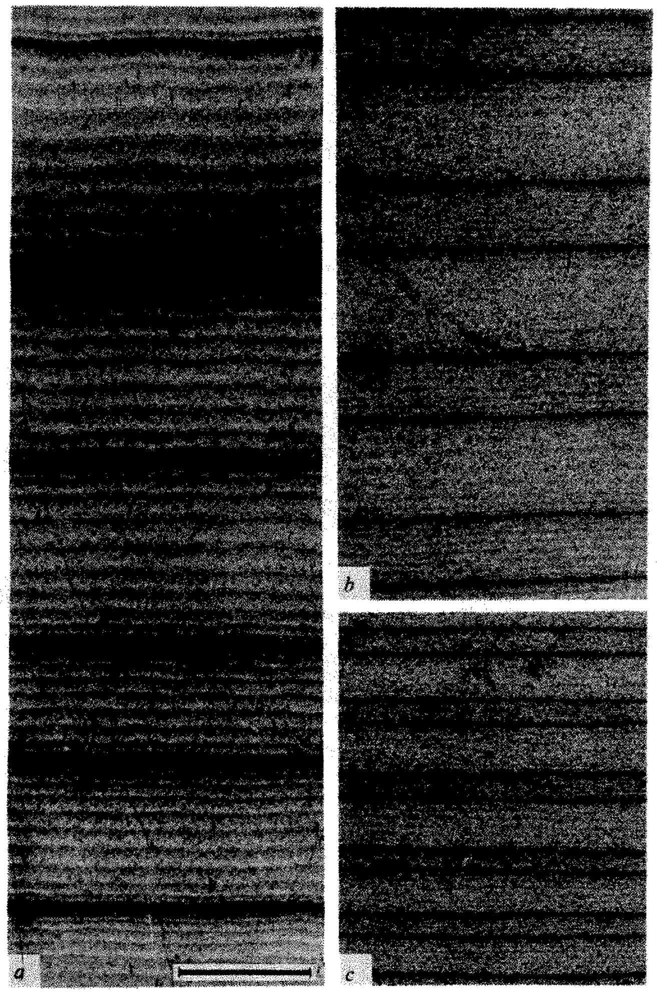


Fig. 1 Cyclic lamination of the Elatina Formation (late Precambrian) at Pichi Richi Pass, South Australia. Dark clayey bands spaced 2–16 mm apart bound cycles of pale, graded laminae of silt to very fine sand grade. *a*, Cycles 9 (base) to 13 (top) of the 16 cycles listed in Table 1. *b* and *c*, 'double cycles' resulting from alternate thick and thin successive cycles. Scale bar, 1 cm.

example, Fig. 2*a*, *b*, *d*) and five exhibit negative skewness (Fig. 2*c*). Overall, the 16 cycles tend to positive skewness (Table 1, column 3).

A 'double cycle' of alternate thick and thin cycles is common (Fig. 1*b*, *c*). Such double cycles are most conspicuous where individual cycles are relatively thin, making accurate counting of clastic laminae difficult. From the laminae counts for thicker cycles, which also exhibit double cycles, it may be concluded, however, that double cycles each contain on average 22–23 clastic laminae. The relative paleness of certain dark bands produces another type of double cycle (for example, cycles 9 and 10 in Fig. 1*a*, which together contain 23 clastic laminae).

Sequences of cycle thicknesses measured between successive dark bands seen on joint-faces and sawcuts perpendicular to the bedding reveal relatively long-term periodicity. Two consecutive sequences containing a total of 157 cycles ($\times 11.2 \simeq 1{,}760$ clastic laminae) display several regularities (see Fig. 3):

(1) The double cycle is prominent through the alternation of thick and thin successive cycles and, where such a pattern may be absent, the presence of paler dark bands between alternate cycles.

(2) Conspicuous long-term variations in cycle thickness are similar in the two sequences. Significant minima occur every 12–14 cycles (~135–155 clastic laminae) and deepest minima every 25–27 cycles (~280–300 laminae); significant maxima occur at intervals of 7–19 cycles (~80–215 laminae) and greatest maxima every 26–27 cycles (~290–300 laminae). The overall pattern suggests two dominant mean periods, one near 26 cycles (~290 laminae) and the other of 13 cycles (~145 laminae). The two periods are in the ratio 2:1 and hence represent a fundamental and a second harmonic. The periods of ~145 and ~290 laminae are conspicuous at other stratigraphic levels and seem to characterize the laminated facies of the Elatina Formation.

(3) Paler dark bands between alternate cycles, best seen on sawcuts, occur in similarly-placed groups within three successive major oscillations (Fig. 3*a*). The stratigraphic occurrence of such paler bands is apparently an integral part of the long-term periodicity identified.

Spectral analysis confirmed the visually-evident periods and revealed also a weaker period of ~90 laminae. Fourier transform of the precisely measured sequence (Fig. 3*a*) identified statistically significant periods (those that can oscillate about 10 times or more in a data sequence[22]) of 22.4 laminae (relative power 2.51) and 89.6 laminae (0.44). For the full sequence (Fig. 3*a*, *b* combined) significant periods identified by Fourier transform are 22.4 laminae (4.89), 89.6 laminae (0.51) and 145.6 laminae (1.37). A period near 285 laminae (1.8) present in the spectrum for the full sequence is in close agreement with the visually-evident period of ~290 laminae. Fourier series analysis of the same sequences confirmed the presence of strong periods near 145 laminae (mean relative power 2.46) and near 290 laminae (4.12).

Late Precambrian climatic cycles

Sedimentary features exhibited by the Elatina Formation—cyclic graded lamination, sporadic graded beds several centimetres thick, cross-laminated sandstones, and rare glacial erratics[23,24]—together suggest deposition mainly by density- and traction-currents in a proglacial lake.

Density currents are of great importance in glaciolacustrine sedimentation[25–29]. During the spring and summer thaw, glacial meltwaters carry abundant suspended matter across the outwash plain to the proglacial lake. There the cold, turbid meltwaters form a density underflow which spreads a thin, commonly graded bed of silt and fine sand widely over the lake floor. Suspended clayey sediment may later settle from quiet waters during the winter when the lake may freeze over, or

Table 1 Data for 16 successive cycles of the Elatina Formation (cycles defined by dark bands)

Cycle	No. of clastic laminae*	Position of thickest lamina from cycle base*
16 (top)	12	4
15	10	6
14	9	4
13	11	6–7
12	14	5
11	11	7
10	11	4
9	12	5
8	13	4
7	14	4
6	11	5
5	12	6
4	12	5
3	9	5–6
2	9	5
1 (base)	10	4–5
Total	180	
Mean (±s.d.)	11.3 ± 1.6	5.0 ± 1.0

* Data obtained through the photographic enlargement and microscopic study of thin sections.

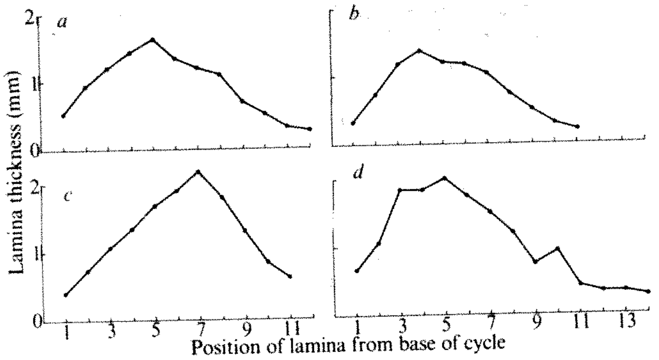


Fig. 2 Variation in thickness of clastic laminae within four cycles shown in Fig. 1*a*. *a*, Cycle 9; *b*, cycle 10; *c*, cycle 11; *d*, cycle 12. Thicknesses of laminae were measured to an accuracy of ~5–10% from 3 to 5 times photographic enlargements of thin sections.

through overturn of lake waters. Clayey winter layers are not always present, however, as summer (seasonal) influx of suspended matter into an oligotrophic lake whose waters are unstratified during influx or do not undergo seasonal overturn can produce graded laminae lacking a clayey uppermost layer[29]. Where such graded beds or coarse-fine couplets form a regular or rhythmic sequence they are usually interpreted as annual deposits, or varves (*sensu lato*)[30,31]. Spasmodic summer storms and dam bursts may also generate stream flows and density currents, resulting in the deposition of several random layers for that year. Glacial varves range in thickness from tens of centimetres to <1 mm, some being too thin to be measured[31]. The thickness of clastic (summer) varves at a particular site may reflect the volume of meltwaters and their suspended load entering the lake; thick clastic varves therefore commonly correspond to warm summers and thin clastic varves to cold summers[32].

The clastic laminae of the Elatina Formation accordingly are ascribed to deposition by density currents probably in a proglacial lake whose waters normally were not subject to winter freezing or overturn. The non-random variation in thickness of such laminae (Fig. 2) indicates a non-random fluctuation also in the volume of successive meltwaters and their suspended load entering the lake. Such regularity of hydrologic and sedimentary processes argues forcibly that the clastic laminae are not the result of random density currents but are indeed annual increments, or varves. Internal layers attributable to spasmodic density currents are uncommon.

As the thickness of clastic varves at a particular site may record relative warmth of summers[32], the regular variation in thickness of varves of the Elatina Formation suggests the cyclic waxing and waning of mean summer temperatures with a mean period near 11 yr. It follows that the dark clayey bands bounding each cycle of ~11 summer laminae were deposited during the coldest portion of the climatic spectrum. The dark bands therefore may be attributed to the deposition of suspended fines during severe winters when the lake froze over or major overturn of cold surface waters occurred. The paler bands suggest a succession of less severe winters at the close of a cycle with limited freezing or overturn. A winter origin for the dark bands implies that the duration of each cycle is recorded by the number of summer clastic laminae.

I conclude therefore that the laminae of the Elatina Formation record a remarkably cyclic climate, with a basic period near

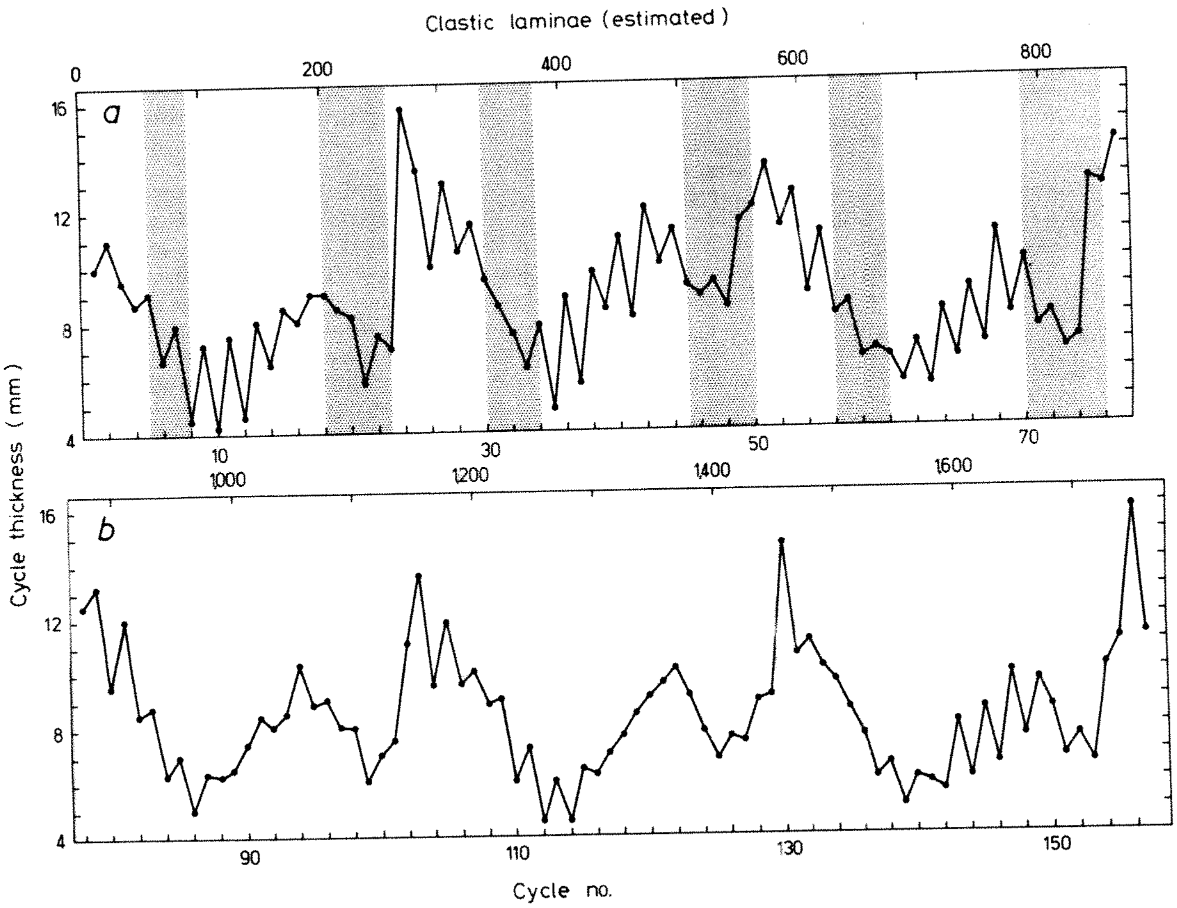


Fig. 3 Long-term trends in cycle thickness for two consecutive vertical sections measured between successive dark bands in the Elatina Formation. The mean cycle is taken as 11.2 clastic laminae. Cycle numbers shown increase up-sequence. *a*, Sawcut; shaded intervals indicate paler darl bands between alternate cycles. *b*, Measured in the field; section immediately overlies that shown in *a*.

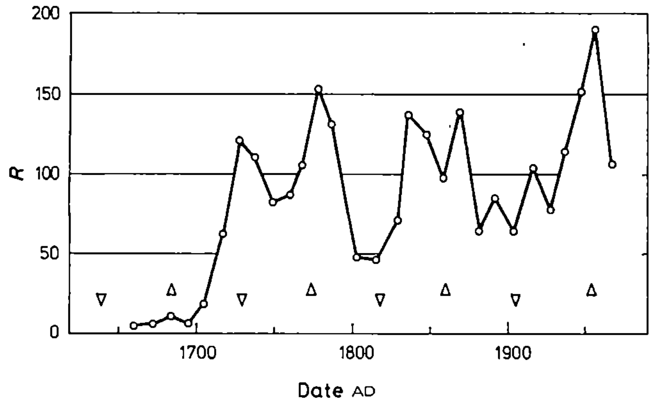


Fig. 4 Annual mean sunspot number R at maxima of the 11-yr solar cycle, from AD 1645 to the present. Triangles mark extrema of the 80–90-yr 'Gleissberg cycle' (from ref. 58).

11 yr, which prevailed in southern Australia during the late Precambrian Marinoan glaciation. During such cycles, summers gradually became warmer over the first 4–7 yr with accompanying increase in the volume and suspended load of glacial meltwaters. The climate then progressively deteriorated over 4–10 yr and the volume and load of meltwaters decreased until at the close of each cycle frigid conditions prevailed. The common alternation in cycle thickness to give a period of ~22 yr (Fig. 1*b*, *c*) suggests the attainment of alternately high and low summer temperatures during successive cycles. Moreover, the interpretation of varve and ~11-yr cycle thicknesses as measures of relative summer temperature implies that long-term variations in cycle thickness (Fig. 3) in effect record long-term fluctuations in relative summer temperature. Hence long-term climatic periods near 90, 145 and 290 yr are revealed, with significant temperature-minima occurring every 135–155 yr and deepest minima every 280–300 yr. Less severe winter end-cycle conditions, as recorded by paler dark bands, appear confined to times of lesser temperature-minima and to intervals immediately preceding the deepest temperature-minima (Fig. 3*a*). Importantly, the climatic-temperature fluctuations were of regional significance, as they apparently affected both the summer meltwater discharges from a continental-type ice sheet and the winter freezing or overturn of an extensive proglacial lake situated ~100 km to the east.

Cause of the climatic cyclicity

The climatic cycles identified may reflect an endogenetic mechanism such as glacial surging, or an exogenetic, solar influence. The idea of climatic control by glacial surging encounters major difficulties:

(1) Surging glaciers[33-35] undergo sudden advance after years of stagnation or very slow retreat. Medvezhii in the Pamirs, for example, has a surge period of 10–14 yr and a surge duration of only 0.2 yr (ref. 33), and mountain glaciers in western North America that have surge periods of 19–20 yr have active phases of only 2–3 yr (ref. 34). The ~11-yr cycles of the Elatina Formation indicate, in contrast, a gradual increase in meltwater discharge over 4–7 yr followed by a waning of such discharge over 4–10 yr. The structure of the ~22-yr double cycles—alternate thick and thin (warmer and cooler) ~11-yr cycles—also seems unrelated to the pattern of glacial surging.

(2) Surging glaciers that have short surge-periods of ≤20 yr typically are relatively small, steep mountain-glaciers that fall through ~1,000–2,000 m and undergo maximum advances of only ~1–5 km (refs 33, 34). In contrast, the regional palaeogeography of the Elatina Formation and correlative shelf units indicates a continental-type ice sheet on a low-relief cratonic area discharging meltwaters to a lake some 100 km distant.

(3) The cycles in the Elatina Formation suggest regional climatic-temperature changes that affected both the ice sheet and lake. A small glacier that periodically advanced only a few kilometres is an unlikely cause of such widespread climatic changes. Indeed, glacial surges are unrelated to climatic fluctuations[34,35].

(4) A surging glacier maintains a reasonably constant surge period[34,35]. The varve-cycles of the Elatina Formation, in contrast, exhibit several strong periods of between ~11 and ~290 yr.

The surging of a glacier therefore seems an improbable control of the climatic cyclicity recorded by the Elatina Formation. Hence I shall examine the Elatina record in detail for evidence of possible solar influence.

Table 2 compares data for 16 successive varve-cycles with two centuries of sunspot data. General correspondence exists between all respective means and ranges for the two sets of data. Moreover, the 16 cycles, which were chosen for study because of their large thickness, tend to positive skewness with respect to varve thickness (Table 1). Interestingly, sunspot cycles of large sunspot maxima on average show positive skewness, the time of rise to maximum sunspot number generally being less than the time of fall to sunspot minimum[36]. Another relevant feature of the Elatina record is the common alternation of thick and thin ~11-yr cycles to produce ~22-yr double cycles. As noted[36,37], sunspot activity between AD 1840 and 1940 exhibited a systematic alternation of high and low maxima for successive 11-yr cycles; the 'zig-zag' pattern of sunspot maxima (Fig. 4) finds similarity in plots of varve-cycle thickness (Fig. 3).

Modern analytical techniques have identified a period near 90 yr in sunspot data since about AD 1700 (refs 2, 38–40), thus confirming the ~80–90 yr 'Gleissberg cycle' (Fig. 4). The period of ~90 yr indicated by Fourier transform of Elatina Formation varve-cycle thicknesses therefore accords with an established sunspot period, but the two strong periods of ~145 and ~290 yr find no apparent counterparts in the observed solar record. The available record of the sunspot number covers less than 300 yr, however, and proposed periods exceeding, say, 90 yr must be regarded with utmost caution[40]. In comparison, the long-term changes in ~11-yr varve-cycle thickness (Fig. 3) cover a total interval of ~1,760 yr.

Study of ^{14}C levels in tree rings suggests four important minima in solar activity since about AD 1000 (Maunder, AD 1654–1714; Spörer, 1416–1534; Wolf, 1282–1342; and a minimum near AD 1040), whose midpoints are spaced respectively 209, 163 and 272 yr apart[41]. The spacing of these inferred solar minima, although irregular, is of the same order as that suggested here for late Precambrian temperature-minima (135–155 and 280–300 yr). Moreoever, one or both of the mean

Table 2 Comparison of data for varve cycles in the Elatina Formation and for recent sunspot cycles

	Mean ±s.d. (yr)	Range (yr)
Data for 16 successive Elatina Formation cycles defined by dark bands (180 yr) (see Table 1)		
Cycle period	11.3±1.6	9–14
Period between thickest varves in successive cycles	11.2±2.4	8–15
Time of rise to maximum varve-thickness in cycle	5.0±1.0	4–7
Time of fall to end of cycle	6.3±2.0	4–10
Sunspot data for 203 yr (19 successive cycles, AD 1755–1958) (ref. 59, p 18):		
Period between minima	11.1±1.3	9.0–13.6
Period between maxima	10.9±2.3	7.3–17.1
Time of rise to maximum	4.5±1.2	2.9–6.9
Time of fall to minimum	6.6±1.4	4.0–10.2

climatic periods of ~145 and ~290 yr recorded by the Elatina Formation accord with: (1) the two dominant climatic periods of ~155 and 280 ± 10 yr found in oxygen and hydrogen stable-isotope ratios measured in 1,800 yr of Japanese cedar rings[42]; (2) the double period of 284–303 yr and apparent second harmonic of 142–151 yr determined in tree-growth from 274 BC to AD 1914 (ref. 43); (3) the period of 290 yr obtained for δD values of bristlecone-pine rings between AD 970 and 1970 (ref. 7); and (4) the period of 286 yr found through ^{14}C measurements on tree rings between 2,000 and 6,000 yr BP (ref. 44).

The ~11-, ~22- and ~90-yr climatic cycles recorded by the Elatina Formation are, therefore, in good agreement with modern sunspot periods, and the cycles of ~145 and ~290 yr with long-term solar and climatic periods as indicated by tree-ring studies. Accordingly, there may be a connection between varve cyclicity and solar variability. Accepting such a relation, the data in Table 2 imply that varve thickness, and therefore meltwater discharge, varied directly with sunspot activity. Maximum summer temperatures thus occurred at sunspot maxima, and the coldest conditions at sunspot minima. By this interpretation, Fig. 3 illustrates long-term trends not only in relative summer temperature but also in relative maximum sunspot number. So delicately balanced were the climatic conditions that only at or near sunspot minima did winter freezing or major overturn of lake waters occur, thus giving the climatic cyclicity such a clear imprint in the sedimentary record.

Discussion

If the climatic cyclicity identified indeed reflects solar variability, then solar control of climate must have been much greater during deposition of the Elatina Formation than is evident today. The intensity of the geomagnetic field may modulate weather and climate[45–47], and palaeointensity variation plausibly may account for such apparent enhanced solar control of climate in the geological past.

The incidence of solar flares and the overall level of flare particle emission, comprising mainly protons, vary directly with the 11-yr solar cycle[2,44,48]. Such corpuscular radiation is normally deflected at the bow shock of the magnetosphere, but some eventually may reach the Earth's atmosphere near the magnetic poles[49,50]. If the dipole component of the geomagnetic field were to weaken or disappear over a few thousand years during a polarity reversal, the shielding effect of the magnetosphere would be much reduced or eliminated; solar corpuscular radiation could then penetrate the atmosphere to lower levels and low magnetic-latitudes and blanket the Earth completely[51–53]. At such times any influence that the sunspot cycle may exert on weather and climate through solar corpuscular radiation[2] would be greatly magnified. The clarity of the varve cycles of the Elatina Formation therefore may be due ultimately to a weakening of the geomagnetic field during a polarity reversal. An alternative explanation is provided by the observation that during the latest Precambrian and early Palaeozoic ~700–400 Myr ago the geomagnetic field apparently was much weaker than that of today[54,55]; a palaeointensity near 10% of the present value is suggested at about the time the Elatina Formation was deposited[54]. This seeming weakness of the late Precambrian geomagnetic field, if confirmed by further studies, might explain the apparently enhanced solar influence on climate at that time.

The empirical evidence presented here suggests that the 11- and 22-yr and longer solar cycles are stable features of the Sun over hundreds of millions of years. Such apparent very long-term stability of solar processes carries important implications for solar physics and models of the solar activity cycle[56]. Moreover, it may allow the prediction of future solar (and, perhaps, climatic) variations from the detailed study of varve cycles where a Sun–climate connection can be shown. The apparent constancy of solar periods over geologically significant time-intervals also has important implications for the long-term constancy of planetary orbital periods[16] and of the gravitational parameter G (ref. 57, p. 341).

The apparent imprint of sunspot periods in Precambrian varved rocks therefore has relevance for solar physics and several areas of planetary science. Ancient annually-layered rocks accordingly should be seen as potential solar observatories that may extend our understanding of solar processes and solar–planetary relations.

I thank P. G. Haslett, M. Kenyon and O. W. Polatayko for valuable assistance, and A. B. Pittock, C. M. Stacey and A. D. Stewart for helpful comments.

Received 22 December 1980; accepted 16 April 1981.

1. Meadows, A. J. *Nature* **256,** 95–97 (1975).
2. Herman, J. R. & Goldberg, R. A. *Sun, Weather, and Climate* (NASA, Washington, 1978).
3. Scherrer, P. H. *Rev. Geophys. Space Phys.* **17,** 724–731 (1979).
4. Lamb, H. H. *Climate: Present, Past and Future* Vol. 1, 440–464 (Methuen, London, 1972).
5. King, J. W. *Astronaut. Aeronaut.* **13**(4), 10–19 (1975).
6. Mitchell, J. M. Jr, Stockton, C. W. & Meko, D. M. in *Solar–Terrestrial Influences on Weather and Climate* (eds McCormac B. M. & Seliga, T. A.) 125–143 (Reidel, Dordrecht, 1978).
7. Epstein, S. & Yapp, C. J. *Earth planet. Sci. Lett.* **30,** 252–261 (1976).
8. Gerety, E. J., Wallace, J. M. & Zerefos, C. S. *J. atmos. Sci.* **34,** 673–678 (1977).
9. Mass, C. & Schneider, S. H. *J. atmos. Sci.* **34,** 1995–2004 (1977).
10. Pittock, A. B. *Rev. Geophys. Space Phys.* **16,** 400–420 (1978).
11. Robock, A. *J. atmos. Sci.* **35,** 1111–1122 (1978).
12. Bradley, W. H. *Prof. Pap. U.S. geol. Surv.* **158,** 87–110 (1929).
13. Anderson, R. Y. *Ann. N.Y. Acad. Sci.* **95,** 424–439 (1961).
14. Richter-Bernburg, G. in *Problems in Palaeoclimatology* (ed. Nairn, A. E. M.) 510–521 (Interscience, London, 1964).
15. Duff, P. McL. D., Hallam, A. & Walton, E. K. *Cyclic Sedimentation* (Elsevier, Amsterdam, 1967).
16. Trendall, A. F. *Econ. Geol.* **68,** 1089–1097 (1973).
17. Preiss, W. V. *Dept Mines Energy, S. Aust., Rep. Bk.* **79/24** (1979); in *Precambrian Geology of the Southern Continents* (ed. Hunter, D. R.) (Elsevier, Amsterdam, in the press).
18. Dalgarno, C. R., Johnson, J. E., Forbes, B. G. & Thomson, B. P. *Port Augusta 1:250,000 Geol. map sheet* (Geological Survey South Australia, Adelaide, 1968).
19. Coats, R. P. *Dept Mines Energy, S. Aust., Rep. Bk.* **79/138** (1979).
20. Coats, R. P. & Preiss, W. V. *Precambr. Res.* **13,** 181–208 (1980).
21. McWilliams, M. O. & McElhinny, M. W. *J. Geol.* **88,** 1–26 (1980).
22. Blackman, R. B. & Tukey, J. W. *The Measurement of Power Spectra* (Dover, New York, 1958).
23. Dalgarno, R. & Johnson, J. E. *Q. Geol. Not. geol. Surv. S. Aust.* **11,** 3–4 (1964).
24. Thomson, B. P. in *Handbook of South Australian Geology* (ed. Parkin, L. W.) 49–83 (Geological Survey of South Australia, Adelaide, 1969).
25. Embleton, C. & King, C. A. M. *Glacial Geomorphology* (Arnold, London, 1975).
26. Jopling, A. V. & McDonald, B. C. (eds) *Glaciofluvial and Glaciolacustrine Sedimentation* (SEPM Spec. Publ. 23, Tulsa, Oklahoma, 1975).
27. Friedman, G. M. & Sanders, J. E. *Principles of Sedimentology* (Wiley, New York, 1978).
28. Schlüchter, C. (ed.) *Moraines and Varves* (Balkema, Rotterdam, 1979).
29. Sturm, M. in *Moraines and Varves* (ed. Schlüchter, C.) 281–285 (Balkema, Rotterdam, 1979).
30. Gary, M., McAfee, R. & Wolf, C. L. (eds) *Glossary of Geology* (American Geological Institute, Washington, 1972).
31. Mörner, N.-A. in *The Encyclopedia of Sedimentology* (eds Fairbridge, R. W. & Bourgeois, J.) 841–843 (Dowden, Hutchinson & Ross, Stroudsburg, 1978).
32. Schove, D. J. in *Moraines and Varves* (ed. Schlüchter, C.) 319–323 (Balkema, Rotterdam, 1979); *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **25,** 209–233 (1978).
33. Budd, W. & McInnes, B. in *Climatic Change and Variability. A Southern Perspective* (eds Pittock, A. B., Frakes, L. A., Jenssen, D., Peterson, J. A. & Zillman, J. W.) 228–234 (Cambridge University Press, 1978).
34. Meier, M. F. & Post, A. *Can. J. Earth Sci.* **6,** 807–817 (1969).
35. Meier, M. F. *Can. J. Earth Sci.* **6,** 987–989 (1969).
36. Bray, R. J. & Loughhead, R. E. *Sunspots* (Chapman & Hall, London, 1964).
37. Dodson, H. W. & Hedeman, E. R. in *Solar-Terrestrial Physics/1970: Part I, The Sun* (ed. de Jager, C.) 151–172 (Reidel, Dordrecht, 1972).
38. Cohen, T. J. & Lintz, P. R. *Nature* **250,** 398–400 (1974).
39. Lomb, N. R. & Andersen, A. P. *Mon. Not. R. astr. Soc.* **190,** 723–732 (1980).
40. Wittman, A. *Astr. Astrophys.* **66,** 93–97 (1978).
41. Stuiver, M. & Quay, P. D. *Science* **207,** 11–19 (1980).
42. Libby, L. M. & Pandolfi, L. J. *Nature* **266,** 415–417 (1977).
43. Turner, H. H. *Mon. Not. R. astr. Soc.* **80,** 793–803 (1920).
44. Damon, P. E. in *The Solar Output and its Variation* (ed. White, O. R.) 429–448 (Colorado Associated University Press, Boulder, 1977).
45. Wollin, G., Ericson, D. B. & Ryan, W. B. F. *Earth planet. Sci. Lett.* **12,** 175–183 (1971).
46. Wollin, G., Ryan, W. B. F. & Ericson, D. B. *Earth planet. Sci. Lett.* **41,** 395–397 (1978).
47. King, J. W. *Nature* **247,** 131–134 (1974).
48. Tandberg-Hanssen, E. *Solar Activity* (Blaisdell, Waltham, 1967).
49. Kavanagh, L. D. Jr in *Earth's Magnetospheric Processes* (ed. McCormac, B. M.) 3–15 (Reidel, Dordrecht, 1972).
50. Lanzerotti, L. J. in *The Solar Output and its Variation* (ed. White, O. R.) 383–403 (Colorado Associated University Press, Boulder, 1977).
51. Siscoe, G. L. & Crooker, N. U. *J. Geomagn. Geoelectr.* **28,** 1–9 (1976).
52. Siscoe, G. L., Chen, C.-K. & Harel, M. *J. atmos. terr. Phys.* **38,** 1327–1331 (1976).
53. Gribbin, J. in *Climatic Change* (ed. Gribbin, J.) 150–154 (Cambridge University Press, 1978).
54. Carmichael, C. M. *Earth planet. Sci. Lett.* **3,** 351–354 (1967); in *Palaeogeophysics* (ed. Runcorn, S. K.) 73–77 (Academic, London, 1970).
55. McElhinny, M. W. *Palaeomagnetism and Plate Tectonics* (Cambridge University Press, 1973).
56. Howard, R. & LaBonte, B. J. *Astrophys. J.* **239,** L33-L36 (1980).
57. Williams, G. E. *Benchmark Papers in Geology* 57 (Hutchinson Ross, Stroudsburg, 1981).
58. Eddy, J. A. in *The Solar Output and its Variation* (ed. White, O. R.) 51–71 (Colorado Associated University Press, Boulder, 1977).
59. Waldmeier, M. *The Sunspot-Activity in the Years 1610–1960* (Schulthess, Zürich, 1961).

Hormonal inducibility of rat α_{2u} globulin genes in transfected mouse cells

David T. Kurtz

Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA

Two different genes coding for the hormonally regulated rat liver protein α_{2u} globulin were introduced into mouse Ltk$^-$ cells through co-transfection with the HSV-1 thymidine kinase gene. Three to ten copies of the α_{2u} globulin genes were detected in several tk$^+$ clones, over 50% of which could be induced with dexamethasone, to produce α_{2u} globulin mRNA and protein. This suggests that the information necessary for hormonal response is contained in the DNA fragment used for transfer.

THE ability to introduce defined DNA segments into mammalian cells[1] is potentially an extremely powerful tool for studying the regulation of eukaryotic gene expression, and for identifying specific DNA sequences which are responsible for such regulation. However, the usefulness of this technique depends critically on obtaining the correct, regulated expression of the foreign gene in the recipient cell. Although expression of rabbit globin mRNA[2] and ovalbumin[3] in mouse L cells following gene transfer has been reported, in neither case was any evidence presented for regulation of expression.

α_{2u} Globulin is a male rat liver protein whose synthesis is under complex hormonal control *in vivo*: androgens, glucocorticoids and thyroid hormone induce α_{2u} globulin synthesis, and oestrogens strongly repress the synthesis of this protein[4,5]. I have recently found that α_{2u} globulin is encoded by a multigene family (18–20 genes per haploid complement), clustered on chromosome 5 (ref. 6). Several α_{2u} globulin genes have been isolated from a library of the rat genome cloned in λ phage. Two of these genes were introduced into mouse Ltk$^-$ cells (which are known to contain receptors for glucocorticoid hormones[7]), to determine if these rat genes retain hormonal inducibility in mouse cells in culture.

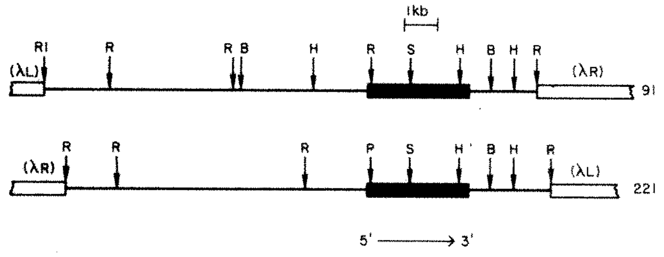


Fig. 1 The α_{2u} globulin genes co-transformed into mouse cells. Clones 91 and 221 are two of the nine α_{2u} globulin clones which have been isolated from the rat library. The coding regions (solid areas) are ~3.2 kilobases (kb) (compared to 1 kb for α_{2u} globulin mRNA) and each gene contains six introns (L. Chow and D.T.K. in preparation). Restriction enzyme sites indicated are: R, *Eco*RI; B, *Bam*HI; S, *Sal*I; H, *Hin*dIII.

Co-transformation of α_{2u} globulin genes into mouse Ltk$^-$ cells

Nine different α_{2u} globulin genes (that is, genes contained on different-sized restriction fragments, and having distinct flanking sequences) have thus far been isolated from a library of the rat genome cloned in Charon 4A. Two of these genes (Fig. 1) were used for co-transformation into mouse Ltk$^-$ cells, using a cloned HSV thymidine kinase gene as a vector[8]. Individual tk$^+$ clones were isolated and grown in mass culture. High molecular weight DNA was extracted from the cell clones, cleaved with *Bam*HI, fractionated on a 0.5% agarose gel and blotted on to nitrocellulose[9]. A cloned cDNA for α_{2u} globulin[10] was labelled by nick translation and hybridized to the filters. it was found (Fig. 2) that the various Ltk$^+$ clones had incorporated 3–10 copies of the α_{2u} globulin genes.

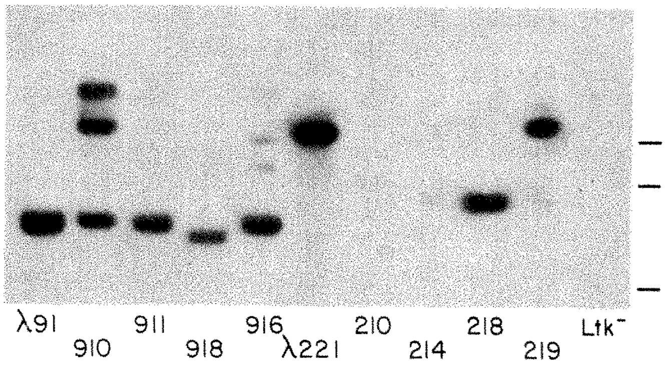


Fig. 2 Presence of α_{2u} globulin gene sequences in mouse L-cell clones. L cells were transformed using the calcium phosphate technique[1] with 20 ng HSV tk gene, plus 500 ng of λ α_{2u} globulin recombinant plus 20 μg carrier mouse Ltk$^-$ DNA per 10^6 cells. Individual tk$^+$ cell clones were isolated, DNA was extracted and cleaved with *Bam*HI. 20 μg were run on a 0.5% agarose gel, blotted on to nitrocellulose and probed with labelled α_{2u} globulin cDNA. Lanes marked 'λ91' and 'λ221' contain 500 pg of λ–α_{2u} recombinants cleaved with *Bam*HI. (This corresponds to ~5 gene equivalents per 20 μg cell DNA.) 910, 911, 916 and 918 are clones which are transformed with λ91, and 210, 218 and 219 were transformed with clone 221.

Introduction of α_{2u} globulin mRNA in response to dexamethasone

L-cell clones which were found to have incorporated the α_{2u} globulin genes were treated with dexamethasone (5×10^{-7} M) for 48 h. Messenger RNA was extracted from the induced and uninduced cells, fractionated on formaldehyde–agarose gels, blotted on to nitrocellulose[11] and probed with nick-translated α_{2u} globulin cDNA. α_{2u} Globulin mRNA was produced in these L cells in response to dexamethasone (Fig. 3). In several of the clones, a constitutive level of α_{2u} globulin mRNA was detectable even in the absence of dexamethasone and the mRNA level in the clones then rose in response to the hormone. No hybridization was found with mRNA from untransformed L cells with or without hormone. One clone (210) was found to produce a mRNA species in response to dexamethasone which hybridized to α_{2u} globulin cDNA, but was considerably larger than authentic α_{2u} globulin message. The nature of this heterogeneity is being investigated.

The induced level of α_{2u} globulin mRNA in these L-cell clones was determined by solution hydridization to single-stranded cDNA (Fig. 4). In a typical clone (918), the level was

0028-0836/81/260629—03$01.00

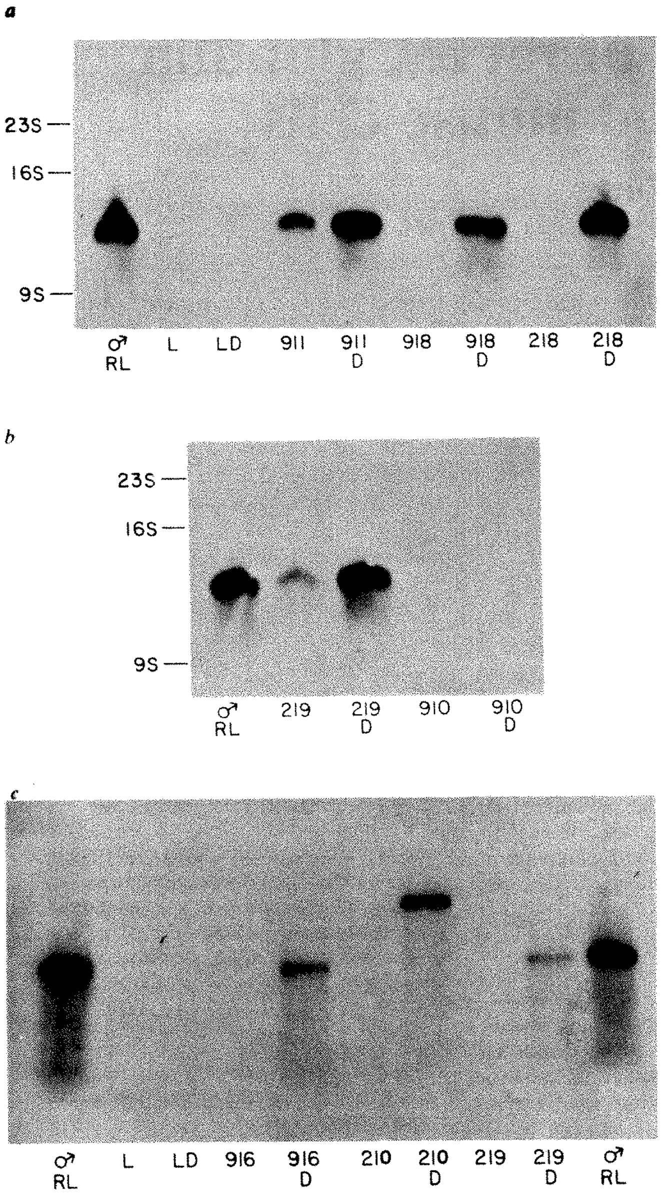


Fig. 3 α_{2u} Globulin mRNA produced in L cells in response to dexamethasone. L-cell clones containing the α_{2u} globulin genes were treated with 5×10^{-7} M dexamethasone for 48 h, and cytoplasmic mRNA was extracted as described previously[10]. The mRNA was fractionated on formaldehyde agarose gels, blotted on to nitrocellulose and probed with the labelled α_{2u} globulin cDNA clone. Lanes marked 'RL' contain 0.1 μg of male rat liver mRNA; all other lanes contain 10 μg of cell mRNA. (Lanes marked 'D' contain mRNA from hormonally induced cells). L, mRNA from untransformed Ltk$^-$ cells. Note that clone 910 produces no detectable α_{2u} globulin mRNA, although it contains several copies of the gene.

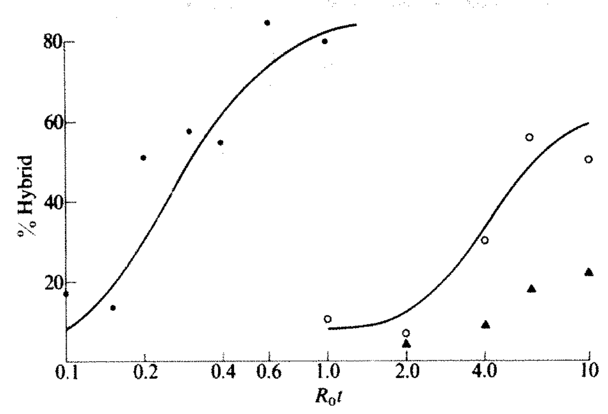


Fig. 4 The cDNA portion of pα176, the pBR recombinant containing α_{2u} globulin cDNA[10], was excised and nick translated (without DNase I) to a specific activity of 5×10^7 c.p.m. per μg. The double-stranded cDNA (1 μg) was denatured and hybridized to 500 μg male rat liver mRNA in 80% formamide, 50 mM PIPES pH 7.4, 5 mM EDTA and 0.4 M NaCl at 55 °C for 2 h. RNA–DNA hybrids were isolated using hydroxylapatite chromatography[4]. The RNA was destroyed by treatment with alkali and the labelled single-stranded cDNA isolated after passage over Sephadex G–50. This single-stranded cDNA (average size ~300 bp) was annealed with male rat liver mRNA (●), or mRNA from transformed cell clone 918, control (▲) or dexamethasone-treated (○). Hybrid formation was determined using S_1 nuclease.

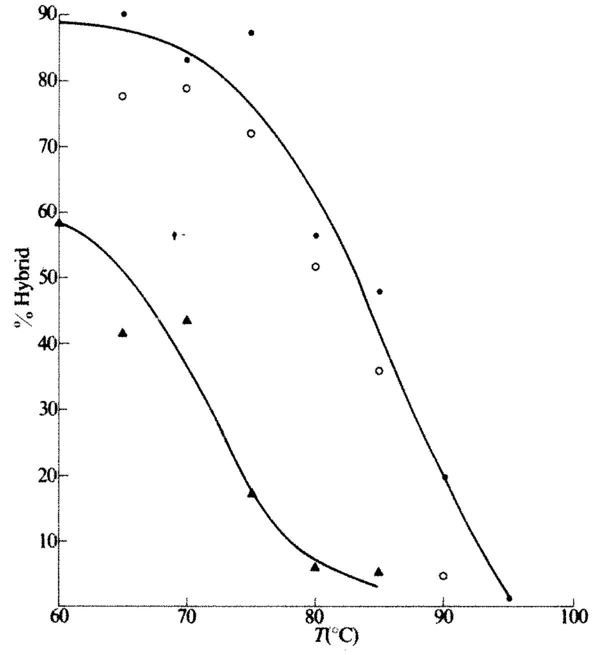


Fig. 5 Thermal denaturation of hybrids between α_{2u} globulin cDNA and MUP mRNA. Hybrids between α_{2u} globulin cDNA and rat liver (●), mouse liver (▲) or clone 918 (○) mRNA were allowed to form at 60 °C. The temperature was raised in 5 °C increments, the tubes were allowed to equilibrate for 10 min, and hybrids were assayed using S_1 nuclease.

~200 copies per cell, that is, 1/10th the level found in adult male rat liver[5].

The cDNA for α_{2u} globulin shows slight cross-hybridization with the mRNA for MUP, the male mouse liver protein which corresponds to rat α_{2u} globulin[10,12]. Therefore, it was important to confirm that the induced mRNA species in these mouse L-cell clones was indeed α_{2u} globulin, and not MUP mRNA. Hybrids formed between α_{2u} globulin cDNA and MUP mRNA melt at 69–70 °C (Fig. 5), compared with a T_m of ~86 °C for α_{2u} globulin mRNA. The T_m of the hybrid between α_{2u} globulin cDNA and the induced mRNA species in clone 918 was found to be identical to that for authentic rat liver α_{2u} globulin mRNA (Fig. 5).

Induction of α_{2u} globulin protein in L-cell clones

It was next of interest to determine if α_{2u} globulin protein was produced and secreted by the cells which contain α_{2u} globulin mRNA. The clones were treated with dexamethasone for 36 h and labelled with ^{35}S-methionine for 16 h. The medium was

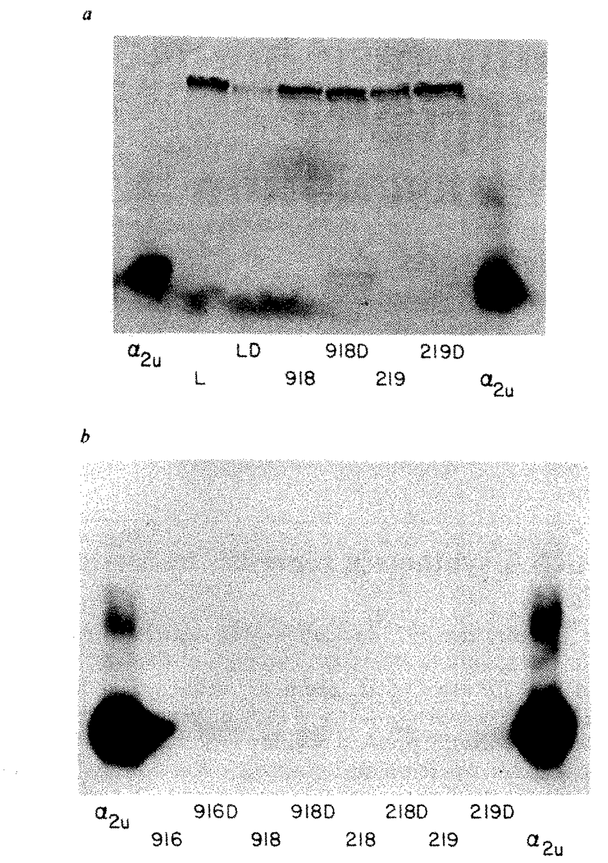


Fig. 6 α_{2u} Globulin protein immunoprecipitated from transformed cell media. Transformed cells were treated with dexamethasone for 36 h, labelled with ^{35}S-methionine (200 μCi per 10^7 cells) for 12 h. The media were collected, immunoprecipitated with rabbit anti-α_{2u} globulin and *S. aureus* (IgG Sorb, New England Enzyme Center). Boiled precipitates were subjected to SDS-gel electrophoresis on 20% polyacrylamide gels[17], and the gels were dried and autoradiographed. Lanes marked 'α_{2u}' contain purified α_{2u} globulin[13] labelled with ^{125}I. Dexamethasone treatment did not enhance overall protein synthesis or secretion in these cells; in most clones, in fact, there was a slight decrease in ^{35}S-methionine incorporation in response to the hormone. *a* Represents a 3-day exposure, *b* a 7-day exposure.

collected and treated with monospecific anti-α_{2u} globulin[13], followed by formalin-fixed *Staphylococcus aureus*. The bacteria were pelleted and washed thoroughly, and the bound protein was released[5] and subjected to SDS-polyacrylamide gel electrophoresis (Fig. 6). ^{35}S-labelled α_{2u} globulin was found in the medium of the L-cell clones which produced α_{2u} globulin mRNA. Again, most of the clones showed a constitutive level of synthesis, which was then further induced in response to hormone. No immunoprecipitated material is found in control (untransformed Ltk$^-$) cells.

Twelve L-cell clones were isolated which contained the entire α_{2u} coding region from λ clone 91; of these, eight were found to produce α_{2u} globulin mRNA and protein in response to hormone. For λ clone 221, four of nine α_{2u}-containing L-cell clones were positive. There seems to be no correlation between α_{2u} gene copy number and the constitutive or induced level of α_{2u} globulin mRNA in the clones.

It is unclear whether the α_{2u} globulin mRNA is being transcribed from one or all of the genes which have been incorporated. Clone 210 produces a mRNA species which hybridizes to α_{2u} globulin cDNA but is ~600 base pairs (bp) larger than authentic α_{2u} globulin message. Further, this mRNA is the only 'α_{2u} globulin' mRNA produced by clone 210. This suggests that, for this particular clone, the transcription may indeed be occurring from only one of the genes, which perhaps has been incorporated into the carrier DNA adjacent to another promoter, or more likely, to another poly(A)-addition site. Preliminary evidence indicates that the large-sized α_{2u} globulin mRNA produced by clones 210 is not a result of incorrect splicing of a correctly transcribed α_{2u} globulin mRNA, because this mRNA does not hybridize to any intron-specific fragments of clone 221.

Discussion

I have shown that rat α_{2u} globulin genes, when introduced into mouse cells, respond to at least one of the hormones which normally induce α_{2u} globulin *in vivo*. (Ltk$^-$ cells possess neither androgen nor oestrogen receptors, and thus, analogous studies concerning the reponse to these hormones must be performed in other tissue culture cells.)

There is no evidence that this hormonal induction of α_{2u} globulin in mouse cells is due to control of the transcription of these genes, although such a model is tempting to consider. Steroid hormones have profound effects on the cytoplasmic half-lifetimes of specific mRNAs in several systems [14,15], and thus the rise in the level of α_{2u} globulin mRNA in these mouse cells could be due to stabilization of the message. Studies are now being done to determine this possibility.

Nonetheless, whether the induction of α_{2u} globulin in these cells is due to transcriptional or post-transcriptional events, it should be possible to identify the DNA sequences in or around the α_{2u} globulin gene which determine the hormonal response, by (1) mutating the α_{2u} globulin genes *in vitro*, reintroducing them into L cells and assessing the effects of the mutations on hormonal response, or (2) ligating specific fragments of the α_{2u} globulin genes to a selectable marker (for example, the HSV tk gene) in an attempt to render such genes hormonally inducible.

Similar studies to those reported here have found[16,18] that mouse mammary tumour virus proviral DNA, cloned in λ phage and co-transformed into L cells, retains the ability to respond to dexamethasone. Thus, in at least two systems, the information necessary for hormonal response is contained in the DNA fragment used in co-transformation. Whether this indicates that the steroid hormone–receptor complex is actually interacting with the DNA (or chromatin) at a site very near or in the gene is unclear; such questions are under investigation.

I thank Doug Bishop for technical assistance, Mike Wigler and Doug Hanahan for the HSV tk gene, and Nancy Hynes and Berndt Groner for making their data available before publication. This work was supported by NIH grant AM2676B to D.T.K., and by NCI Cancer Center Research grant CA13106 to J. D. Watson.

Received 23 February; accepted 29 April 1981.

1. Wigler, M. *et al. Cell* **16**, 777–785 (1979).
2. Mantei, N., Boll, W. & Weissman, C. *Nature* **281**, 40 (1979).
3. Lai, E. C., Woo, S. L. C., Bordelon-Riser, M. E., Fraser, T. F. & O'Malley, B. *Proc. natn. Acad. Sci. U.S.A.* **77**, 244 (1980).
4. Kurtz, D. T. & Feigelson, P. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4791–4795 (1977).
5. Kurtz, D. T., Chan, K.-M. & Feigelson, P. *Cell* **15**, 743–750 (1978).
6. Kurtz, D. T. *J. molec. appl. Genet.* (in the press).
7. Lippman, M. E. & Thompson, E. B. *J. biol. Chem.* **249**, 2483–2488 (1974).
8. Hanahan, D. Lane, D., Lipsich, L., Wigler, M. & Botchan, M. *Cell* **21**, 127–139 (1980).
9. Southern, E. M. *J. molec. Biol.* **18**, 503–517 (1975).
10. Kurtz, D. T. & Nicodemus, C. F. *Gene* **13**, 145–152 (1981).
11. Thomas, P. S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5201–5205 (1980).
12. Hastie, N. D., Held W. A. & Toole, J. T. *Cell* **17**, 449 (1979).
13. Kurtz, D. T., Sippel, A. E. & Feigelson, P. *Biochemistry* **15**, 1031–1036 (1976).
14. Cox, R. F. *Biochemistry* **16**, 3433 (1977).
15. Guyette, W. A., Matusik, R. J. & Rosen, J. R. *Cell* **17**, 1013 (1979).
16. Hynes, N. E., Kennedy, N., Rahmsdorf, U. & Groner, B. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2038–2042 (1981).
17. Laemmli, U. K. *Nature* **227**, 680–685 (1970).
18. Buetti, E. & Diggelmann, H. *Cell* **23**, 335–346 (1981).

An aminoacyl tRNA synthetase binds to a specific DNA sequence and regulates its gene transcription

Scott D. Putney & Paul Schimmel

Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Alanine tRNA synthetase represses transcription of its own gene by binding specifically to a palindromic sequence which flanks the gene's transcription start site. Transcription repression is greatly enhanced by elevated concentrations of the cognate amino acid. The amino acid effect is caused by direct association of the ligand with the synthetase, which in turn mediates tighter binding to the DNA.

AMINOACYL tRNA synthetases carry out the first step in protein biosynthesis by attaching amino acids to the corresponding tRNAs[1]. The cellular levels of these enzymes are regulated in two different ways. The first, which has been demonstrated for approximately half of the 20 synthetases, is derepression of enzyme levels during restriction of the cognate amino acid[2–6]. The second, termed metabolic regulation, is the dependence of synthetase levels on cellular growth rate; all synthetases studied thus far show this type of regulation[7], as do stable RNA species[8,9] and other proteins involved in protein synthesis.[9–12].

The work presented here deals with the molecular mechanism for the first form of regulation, although, as explained later, it may also pertain to metabolic regulation. We previously reported the cloning and partial sequencing of the gene for *Escherichia coli* Ala-tRNA synthetase[13]. This enzyme has an α_4 structure with a subunit molecular weight of 95,000 (ref. 14) making it the largest bacterial aminoacyl tRNA synthetase[1,14]. We sequenced the 5'-coding region of the structural gene and ~250 nucleotides upstream from the start codon for the polypeptide[13]. *In vitro* transcription studies showed that initiation occurs at −79 (relative to the first base of the polypeptide start codon which is designated as +1) and that, between the transcription initiation point and the polypeptide start, there is no leader peptide-encoding region or attenuator as occurs in numerous amino acid biosynthetic operons[13,15–22]. Thus, if transcription of *ala*S is subject to amino acid-dependent regulation, a different kind of mechanism must be operative. It was with this in mind that we expanded our initial investigations of *in vitro* transcription of this operon.

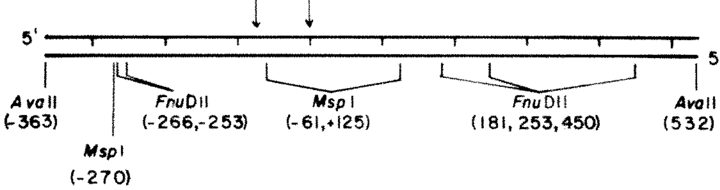


Fig. 1 Schematic diagram of the *Ava*II restriction fragment used for transcription studies. The fragment is 895 base pairs long and contains 532 nucleotides of the structural gene of *ala*S. Position +1 is assigned to the ATG at the beginning of the structural gene, and transcription is initiated with pppA at position −79. The fragment was prepared, as described previously, by purifying the plasmid pSP201 (which contains a 3.7-kilobase section of the vector pBR322 and a large segment of *ala*S) and digesting it with *Ava*II[13]. The resulting mixture was electrophoresed through a 5% (30:1) acrylamide gel and the above fragment recovered by the method of Maxam and Gilbert[58].

Ala-tRNA synthetase represses transcription of *ala*S

Figure 1 shows the 895 nucleotide-long restriction fragment used for transcription, where the two ends of the fragment are located at −363 and +532. From the sequence of the end-labelled primary transcript, it was shown previously that transcription initiates with pppA at position −79 (ref. 13). When this fragment is used as a template for transcription with [α-^{32}P]UTP as the sole labelled ribonucleotide, one major RNA product is synthesized (Fig. 2*A*, lane *a*). The arrow indicates the position of the transcript initiated at position −79. Figure 2*A* also shows the effect of 1–10 μM concentrations of Ala-tRNA synthetase on transcription. As more synthetase is added, there is clearly a progressive decrease in the amount of primary transcript produced.

To show that Ala-tRNA synthetase does not affect promoters other than *ala*S, a 984-base pair restriction fragment containing the promoter for the histidine operon[17] was transcribed in the absence and presence of the synthetase. Transcription of this fragment gives rise to an ~168-base pair leader RNA species which is heterogeneous in length at its 3'-terminus owing to stuttering of the polymerase at the string of uridines in the termination site (R. Freedman, personal communication). Figure 2*B* shows that, at a concentration of Ala-tRNA synthetase which represses transcription of *ala*S (lanes *a*, *b*), there is no effect on transcription from the *his* promoter (lanes *c*, *d*). Moreover, when the fragment containing the *ala*S promoter is mixed with that containing the *his* promoter, transcription proceeds cleanly from both templates (lane *e*) while, in the presence of the synthetase, only *ala*S transcription is repressed (lane *f*).

The effect on transcription of proteins other than Ala-tRNA synthetase was also investigated. Figure 2*C* shows the effect of 2 mg ml^{-1} concentrations of Ala, Ile and Tyr-tRNA synthetases, as well as bovine serum albumin and lysozyme. Because all these proteins have significantly lower molecular weights than Ala-tRNA synthetase, their molar concentrations in the transcription mixture are several-fold higher than that of Ala-tRNA synthetase. Nevertheless, only the alanine-specific synthetase represses transcription of *ala*S.

Collectively, the data in Fig. 2*A*, *B* and *C* establish that transcription of *ala*S is specifically repressed by micromolar concentrations of the cognate synthetase.

Autoregulation of *ala*S transcription depends on alanine concentration

Figure 2*D* shows the effects of 0, 1 and 2 mM alanine on the Ala-tRNA synthetase-mediated repression of *ala*S transcription. Only 1 μM of synthetase was used in the transcription mixture. As the alanine concentration is increased in the

0028–0836/81/260632—04$01.00

presence of synthetase, there is a concomitant decrease in the amount of primary transcript synthesized. Many alanine concentrations were used in other experiments which utilized the synthetase at a 10-fold higher concentration and the results were analogous to those shown in Fig. 2*D*. Control experiments showed that there is no effect of 4 mM alanine in the absence of synthetase nor of 4 mM glycine or serine in the presence of synthetase (data not shown). Although the extent of the alanine effect has not been rigorously quantified, we estimate that, in the presence of synthetase plus 1.5 mM alanine, there is more than a 20-fold decrease in the amount of transcript synthesized relative to that synthesized in the presence of synthetase alone.

To understand more fully the interaction between alanine and Ala-tRNA synthetase, we measured the K_m for alanine in the standard amino acid-dependent ATP–PP_i exchange reaction[23]. In conditions of approximate pyrophosphate and ATP saturation, the K_m for alanine was found to be 1.1 mM at *p*H 7.5 (data not shown). (This value is 10- to 100-fold greater than the amino acid K_ms of other *E. coli* aminoacyl tRNA synthetases[24–33].) Because the concentration of alanine required to affect transcription is comparable with the independently measured K_m, we infer that the site which binds alanine for aminoacylation is the same as that which confers alanine dependence to transcription regulation.

In many of the gels shown in Fig. 2 there is a small amount of *in vitro* synthesized RNAs which migrate above the primary transcript. Previously we estimated that the major species constitutes ~90% of the RNA made from this template[13]. Transcription of these minor species is repressed in the presence of sufficient Ala-tRNA synthetase or synthetase plus alanine. This is probably because binding of the synthetase around the transcription initiation point of −79 (see below) blocks the complete synthesis of these presumably spurious, non-specific transcripts which originate upstream from −79.

Ala-tRNA synthetase binds to a region flanking transcription initiation

Because Ala-tRNA synthetase specifically affects transcription of *ala*S, it seemed plausible that the synthetase directly interacts with the DNA template itself. To test this possibility, DNA

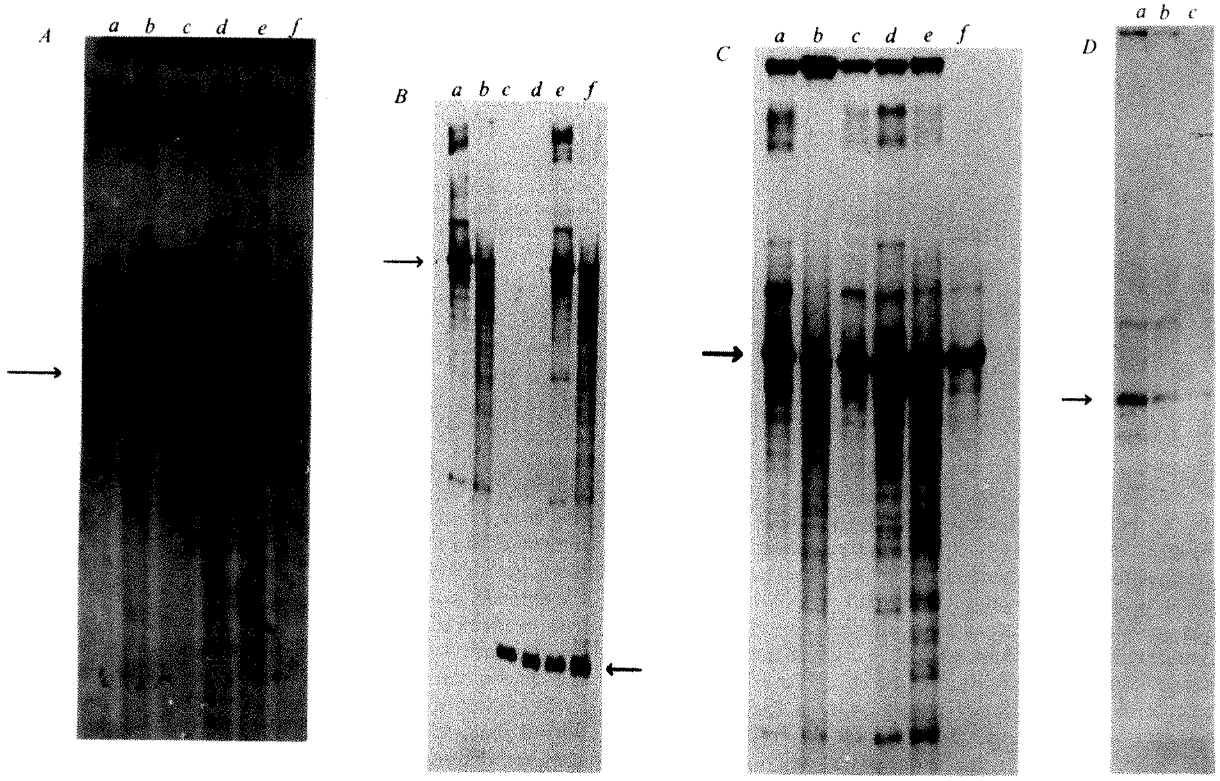
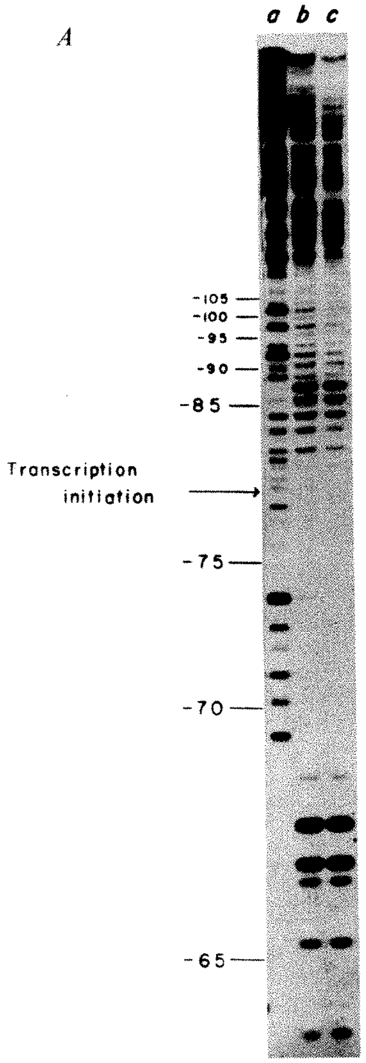


Fig. 2 *A*, the effect of varying concentrations of Ala-tRNA synthetase on the *in vitro* transcription of *ala*S. An autoradiograph of a denaturing polyacrylamide gel through which radioactively labelled products of transcription reactions were electrophoresed is shown. The arrow indicates the position of the ~620-nucleotide *ala*S primary transcript which begins 79 nucleotides before the start of the structural gene (see Fig. 1). The intensity of these bands indicates the amount of transcript made in each of the six reactions. The amount of Ala-tRNA synthetase present is (in μM): *a*, 0; *b*, 1.0; *c*, 2.5; *d*, 5.0; *e*, 7.5; *f*, 10.0. As the concentration of Ala-tRNA synthetase in the transcription reaction increases, the transcription of *ala*S is repressed. (The primary transcript band in lane *a* appears slightly less intense than that in lane *b* because, due to the larger width of lane *a*, the radioactivity covered a larger area.) Transcription reactions were performed at 37 °C in a 10-μl volume and contained 10 mM Tris-HCl *p*H 8.0, 10 mM $MgCl_2$, 0.1 mM EDTA, 150 mM KCl, 1 mM dithiothreitol, 0.5 μg purified RNA polymerase (a gift from Professor Will McClure), 50 μM each of the four ribonucleotide triphosphates (with the only labelled ribonucleotide being [α-^{32}P]UTP (Amersham) at a specific activity of 20 Ci $mmol^{-1}$), 0.1 μg of the *Ava*II fragment and the indicated amount of Ala-tRNA synthetase (purified to homogeneity using the procedure of Putney *et al.*[14]). The template and the synthetase were incubated for 10 min before adding the other ingredients. Reactions continued for 20 min, 40 μl of 100 μg ml^{-1} unfractionated tRNA was then added and the mixture phenol-extracted followed by ether extraction. After ethanol precipitation the pellet was dissolved in 20 mM sodium citrate, *p*H 5.0, 1 mM EDTA with 9 M urea containing tracking dyes, heated to 100 °C for 1 min and electrophoresed on a 6% (20:1) polyacrylamide gel containing 7 M urea. The gel was then autoradiographed. *B*, the effect of Ala-tRNA synthetase on the transcription of *ala*S and on the transcription of a 984-base pair *Hin*dII restriction fragment containing the promoter for the *S. typhimurium* histidine operon[17]. The arrow on the left indicates the position of the *ala*S primary transcript and the arrow on the right the position of the *his* leader transcript. Lanes *a* and *b* display products of transcription reactions which contained the *ala*S fragment (Fig. 1) as the template, the products in *c* and *d* are from reactions which contained the histidine operon template (a gift from Mr Richard Freedman) and the products in *e* and *f* from reactions which contained both templates. The reactions represented by lanes *a*, *c* and *e* had no Ala-tRNA synthetase whereas the reactions for lanes *b*, *d* and *f* had 10 μM Ala-tRNA synthetase. The results show that the presence of Ala-tRNA synthetase represses synthesis of the *ala*S transcript but does not affect synthesis of the *his* leader transcript. *C*, the effect of various proteins on the transcription of *ala*S. The *Ava*II fragment (Fig. 1) was transcribed in the presence of the following purified proteins: *a*, no added protein; *b*, 2 mg ml^{-1} (5 μM) Ala-tRNA synthetase; *c*, 2mg ml^{-1} (18 μM) Ile-tRNA synthetase (a gift from Dr Stephen Koontz); *d*, 2 mg ml^{-1} (21 μM) Tyr-tRNA synthetase (a gift from Dr Rochester Co); *e*, 2 mg ml^{-1} (30 μM) bovine serum albumin (Sigma); and *f*, 2 mg ml^{-1} (140 μM) lysozyme (Sigma). Only when Ala-tRNA synthetase is present during transcription is the amount of the *ala*S transcript diminished. *D*, the effect of alanine on the repression of *ala*S transcription by Ala-tRNA synthetase. The *Ava*II fragment (Fig. 1) was transcribed in the presence of 1 μM Ala-tRNA synthetase and various concentrations of alanine. Alanine concentrations were (in mM): *a*, 0; *b*, 1; *c*, 2. As the concentration of alanine in the transcription reaction increases, the repression of *ala*S transcription by Ala-tRNA synthetase is enhanced. Reactions and product analyses were performed as described in legend to *A*.



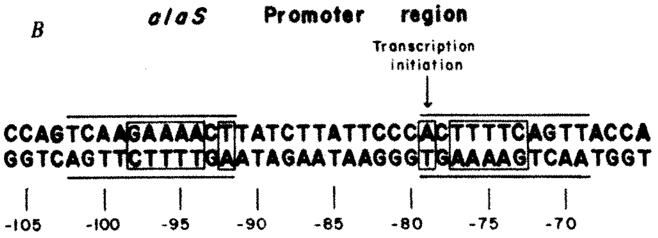

B
alaS Promoter region
Transcription initiation
CCAGTCAAGAAAACTTATCTTATTCCCACTTTTCAGTTACCA
GGTCAGTTCTTTTGAATAGAATAAGGGTGAAAAGTCAATGGT
-105
-100
-95
-90
-85
-80
-75
-70

Fig. 3 *A*, protection of the *ala*S promoter from digestion by DNase I by Ala-tRNA synthetase. The autoradiograph is of a denaturing gel through which were electrophoresed DNase I digestion products of a double-stranded restriction fragment containing the *ala*S promoter. The fragment is generated by the restriction enzymes *Fnu*DII and *Msp*I, and spans bases −253 to −59 (see Fig. 1); it was end-labelled at −59. Lane *a* shows the products in the absence of Ala-tRNA synthetase, and the reactions in lanes *b* and *c* were incubated with 10 and 20 μM synthetase, respectively. To the left are numbers indicating the positions of the nucleotides, and the position of transcription initiation is indicated (Figs 1, 3*B*). These positions were determined by electrophoresing, in parallel, products of sequencing reactions[58] of the end-labelled fragment. In the presence of Ala-tRNA synthetase, cleavage at nucleotides −69 to −79 and −92 to −102 is repressed, indicating that the synthetase binds the DNA at these regions. Digestion reactions were performed at 37 °C in 10-μl volumes and contained 10 mM Tris-HCl (*p*H 8.0), 10 mM $MgCl_2$, 0.1 mM EDTA, 150 mM KCl, 1 mM dithiothreitol and 1 mM $CaCl_2$. 0.012 μg of DNA, which was isolated after labelling at the *Msp*I site at position −59 by the kinase exchange reaction[59] followed by restriction at the *Fnu*DII site at position −253, was incubated in the cocktail for 10 min before adding 0.1 ng DNase I (grade I, Boehringer Mannheim). The reaction was terminated after 15 min by the addition of 50 μl 100 μg ml^{-1} sonicated calf thymus DNA. Treatment was then as described in Fig. 2*A* legend, except that electrophoresis was through a 10% (20:1) polyacrylamide gel containing 7 M urea. *B*, nucleotide sequence of the region of *ala*S which flanks the transcription initiation site and the section protected from DNase I digestion by Ala-tRNA synthetase (see *A*). The nucleotide sequence was determined by Putney *et al.*[13]; numbers below the sequence indicate positions before the start of the structural gene, and transcription is initiated at the A shown by the arrow at position −79. The nucleotides contained between underlines and overbars are those which are protected by the synthetase from DNase I digestion (see *A*), and the boxed nucleotides form the palindrome within the protected regions. This palindrome flanks the transcription initiation site.

5′ GAAAANT 3′
3′ CTTTTNA

footprint experiments[34,35] were carried out with restriction fragments containing the transcription initiation region. In preliminary experiments we demonstrated an interaction of the synthetase with a section roughly between −120 and −60. To define more precisely this interaction, an *Fnu*DII–*Msp*I restriction fragment which runs from −253 to −59 was prepared (see Fig. 1) and 5′-end-labelled with ^{32}P at −59. The results of an experiment with this fragment are shown in Fig. 3*A*. The synthetase clearly protects nucleotides −69 through −79 from DNase I digestion; in addition, nucleotides −92 to −102 are protected. (This is more obvious on a gel on which the identical digest was electrophoresed further to improve the separation of individual fragments in the region of −85 to −110.) No other regions are as unambiguously protected by the presence of the synthetase. In addition, the synthetase enhances cleavage at certain sites, and this phenomenon has also been observed in other DNase protection studies[36].

Figure 3*B* summarizes these results. The nucleotide sequence previously determined for the region −106 to −65 is shown. Transcription initiation occurs at −79 and starting at −91 there is a sequence, TATCTTA, which fulfills the main features of a consensus sequence (Pribnow box) associated with RNA polymerase promoter sites[13,37]. Also shown are the two nucleotide segments which bind to Ala-tRNA synthetase. These two segments flank the promoter sequence and transcription initiation point.

Within each of the protected regions is a palindromic sequence:

5′GAAAANT
CTTTTNA 5′

the centres of which are 19 base pairs apart, corresponding almost exactly to two turns of a DNA B helix. Therefore, the palindromic sequences are perfectly spaced so as to interact, on the same face of the DNA, with a protein having a 2-fold axis of symmetry. Previous intersubunit cross-linking data obtained with dimethyl suberimidate exclude cyclic 4-fold symmetry for the native tetrameric Ala-tRNA synthetase, but are easily explained with a 2-fold symmetry for the protein[14]. Moreover, with a molecular weight of 380,000, the synthetase can easily span the distance of 70–100 Å implied by the results of Fig. 3*B*, assuming a DNA B-helical structure. Alternatively, the palindromic sequences could arrange themselves in a hairpin-type of cloverleaf structure which interacts with the synthetase.

Implications and physiological significance

As mentioned earlier, amino acid-dependent repression of aminoacyl tRNA synthetase synthesis has been demonstrated in several cases, although no analyses have been reported for *E. coli* Ala-tRNA synthetase[2–6]. Based on data of Piperno and Oxender[38] we estimate that the intracellular concentration of alanine is between 1 and 5 mM in exponentially growing *E. coli* K12. This is substantially greater than the pool sizes of most other amino acids and, significantly, the K_m for alanine (for Ala-tRNA synthetase, ~1 mM in our assay conditions) is substantially higher than that reported so far for other synthetases. Thus, to a first approximation, the K_m and the pool size for alanine are comparable and, therefore, amino acid-dependent regulation might take place under these circumstances.

From enzyme assays of Putney *et al.*[13], the intracellular concentration of Ala-tRNA synthetase in cells growing in rich media is 0.4 μM; this corresponds to 250 molecules per cell. This is in close agreement with the values of 300–900 molecules per cell for various other tRNA synthetases[7].

At a concentration of 0.4 μM, there is little if any autogenous repression in the absence of alanine (Fig. 2*A*). Hence, under these circumstances, *ala*S autogenous regulation is mediated entirely by the intracellular alanine concentration. As shown in Fig. 2*D*, effective alanine-dependent repression of transcription occurs *in vitro* at a synthetase concentration of 1 μM and presumably at somewhat lower concentrations also. Because the transcription reaction must be carried out in the presence of ATP, we do not know whether the critical liganded species is the synthetase–amino acid, synthetase–amino acid–ATP, or synthetase–aminoacyl adenylate complex. Note, however, that the K_m for alanine was also determined in the presence of ATP (in the ATP–PP_i exchange reaction), so it can be compared directly with the concentrations of alanine required for repression of transcription. In addition, preliminary results have indicated no additional effects of $tRNA^{Ala}$ (when added with ATP and amino acid) beyond those achieved by the amino acid alone in the transcription mixture.

The qualitative parallels between the interaction of Ala-tRNA synthetase with the *ala*S promoter region and the *lac* repressor–*lac* operator system are striking. In both cases, the protein (Ala-tRNA synthetase or *lac* repressor) binds to a palindromic sequence near the transcription initiation point[34,39]. In both instances, a ligand (amino acid or sugar, respectively) alters the affinity of the protein for the DNA, although the effects are exerted in the opposite way (alanine enhances binding of the synthetase whereas sugar binding to the repressor weakens the interaction). In addition, both proteins are tetramers and the *lac* repressor has 2-fold symmetry[40].

The work done on the amino acid-dependent repression–derepression of aminoacyl tRNA synthetase levels[2–6] shows that synthesis is derepressed in conditions of cognate amino acid starvation, and there are synthetase mutants which display properties consistent with amino acid-dependent autogenous regulation, such as a mutation in His-tRNA synthetase which results in an elevated K_m for histidine[41]. This alteration causes derepression of synthesis at histidine levels which cause repression in the wild-type strain. If histidine-dependent autoregulation is operative for this enzyme,. the elevated K_m would result in a decreased level of histidine-complexed enzyme and, consequently, less stringent repression. (There is some evidence, however, that the effector molecule is $tRNA^{His}$ rather than free histidine[41,42].) Another set of examples involves mutations, located outside the synthetase structural gene, which result in increased enzyme levels. For Ser-, Val- and Leu-tRNA synthetases, these alterations are linked to the respective structural gene and are thought to be mutations in the operator locus[43–46]. Such results could be explained by autogenous control of the type described here if the mutation occurred at the synthetase binding site in the promoter region.

Aminoacyl tRNA synthetases are also subject to metabolic regulation whereby their intracellular concentrations vary with growth rate conditions[2–7] as do other species involved in protein synthesis[8–12]. Neidhardt *et al.* have proposed a mechanism for both amino acid-dependent repression and metabolic control involving just the substrates and products of the aminoacylation reaction[2]. This mechanism would also explain amino acid-dependent repression of aminoacyl tRNA synthetase synthesis. On the other hand, there is evidence that guanosine tetraphosphate is related to aminoacyl tRNA synthetase regulation[47,48], as well as to aminoacyl tRNA synthesis[49,50] and enzyme synthesis from operons which are subject to amino acid control via the attenuator mechanism[51,52]. At present our data only pertain to the mechanism of amino acid-dependent repression; however, with suitable modification, the Neidhardt *et al.* proposal might be harmonized with our data to explain both amino acid-dependent repression and metabolic control, but it is likely that more considerations are involved than have been explored here.

Several ribosomal proteins, the levels of which are under metabolic regulation, are also autogenously controlled[51–57]. However, in these cases, much of the effect seems to be at the translational level (although transcription may also be affected[57]), and no co-repressor has been found as for Ala-tRNA synthetase. We have not investigated the possibility of autoregulation of translation of Ala-tRNA synthetase, but in view of the effective transcriptional control which operates we question the need for (but do not rule out) an additional translation based regulatory mechanism.

This work was supported by NIH grant GM15539. S.P. was a NIH Predoctoral Trainee from 1977 to 1980.

Received 29 December 1980; accepted 30 March 1981.

1. Schimmel, P. R. & Söll, D. *A. Rev. Biochem.* **48,** 601–648 (1979).
2. Neidhardt, F. C., Parker, J. & McKeever, W. G. *A. Rev. Microbiol.* **29,** 215–250 (1975).
3. Brenchley, J. E. & Williams, L. S. *A. Rev. Microbiol.* **29,** 251–274 (1975).
4. La Rossa, R. & Söll, D. in *Transfer RNA* (ed. Altman, S.) 136–167 (MIT Press, Cambridge, Massachusetts, 1978).
5. Umbarger, H. E. in *Transfer RNA: Biological Aspects* (eds Söll, D., Abelson, J. & Schimmel, P.) 453–467 (Cold Spring Harbor Laboratory, New York, 1980).
6. Morgan, S. D. & Söll, D. *Prog. Nucleic Acid Res. molec. Biol.* **21,** 181–207 (1978).
7. Neidhardt, F. C., Bloch, P. L., Pedersen, S. & Reeh, S. *J. Bact.* **129,** 378–387 (1977).
8. Skjold, A. C., Jauarez, H. & Hedgcoth, C. *J. Bact.* **115,** 177–187 (1973).
9. Kjeldgaard, N. O. & Gausing, K. in *Ribosomes* (eds Nomura, M., Tissieres, A. & Lengyel, P.) 369–392 (Cold Spring Harbor Laboratory, New York, 1974).
10. Gordon, J. *Biochemistry* **9,** 912–917 (1970).
11. Pedersen, S., Bloch, P., Reeh, S. & Neidhardt, F. *Cell* **14,** 179–190 (1978).
12. Herendeen, S., Van Bogelen, R. & Neidhardt, F. *J. Bact.* **139,** 185–194 (1979).
13. Putney, S. D., Melendez, D. & Schimmel, P. R. *J. biol. Chem.* **256,** 205–211 (1981).
14. Putney, S. D., Sauer, R. T. & Schimmel, P. R. *J. biol. Chem.* **256,** 198–204 (1981).
15. Bennett, G. N., Schweingruber, M. E., Brown, K. D., Squires, C. & Yanofsky, C. *Proc. natn. Acad. Sci. U.S.A.* **73,** 2351–2355 (1976).
16. Di Nocera, P. P., Blasi, F., Di Lauro, R., Frunzio, R. & Bruni, C. B. *Proc. natn. Acad. Sci. U.S.A.* **75,** 4276–4280 (1978).
17. Barnes, W. M. *Proc. natn. Acad. Sci. U.S.A.* **75,** 4281–4285 (1978).
18. Zurawski, G., Brown, K., Killinglby, D. & Yanofsky, C. *Proc. natn. Acad. Sci. U.S.A.* **75,** 4271–4275 (1978).
19. Gardner, J. F. *Proc. natn. Acad. Sci. U.S.A.* **76,** 1706–1710 (1979).
20. Gemmill, R. M., Wessler, S. R., Keller, E. B. & Calvo, J. M. *Proc. natn. Acad. Sci. U.S.A.* **76,** 4941–4945 (1979).
21. Nargang, F. E., Subrahmanyam, C. S. & Umbarger, H. E. *Proc. natn. Acad. Sci. U.S.A.* **77,** 1823–1827 (1980).
22. Lawther, R. P. & Hatfield, W. *Proc. natn. Acad. Sci. U.S.A.* **77,** 1862–1866 (1980).
23. Calendar, R. & Berg, P. in *Procedures in Nucleic Acids Research* (eds Cantoni, G. L. & Davies, D. R.) 384–399 (Harper & Row, New York, 1966).
24. Calendar, R. & Berg, P. *Biochemistry* **5,** 1690–1695 (1966).
25. Kalousek, F. & Konigsberg, W. *Biochemistry* **13,** 999–1006 (1974).
26. Lemoine, F., Waller, J. P. & Rapenbusch, R. *Eur. J. Biochem.* **4,** 213–221 (1968).
27. Katze, J. R. & Konigsberg, W. *J. biol. Chem.* **245,** 923–930 (1970).
28. Joseph, D. & Meunch, K. *J. biol. Chem.* **246,** 7602–7609 (1971).
29. Hirshfield, I. & Bloemers, H. P. J. *J. biol. Chem.* **244,** 2911–2916 (1969).
30. Rouget, P. & Chapeville, F. *Eur. J. Biochem.* **14,** 498–508 (1970).
31. Yaniv, M. & Gros, F. *J. molec. Biol.* **44,** 1–15 (1969).
32. Cole, F. X. & Schimmel, P. R. *Biochemistry* **9,** 480–489 (1970).
33. Ostrem, D. & Berg, P. *Biochemistry* **13,** 1338–1348 (1974).
34. Galas, D. J. & Schmitz, A. *Nucleic Acids. Res.* **5,** 3157–3170 (1978).
35. Johnson, A. D., Meyer, B. J. & Ptashne, M. *Proc. natn. Acad. Sci. U.S.A.* **76,** 5061–5065 (1979).
36. Schmitz, A. & Galas, D. J. *Nucleic Acids Res.* **6,** 111–137 (1979).
37. Rosenberg, M. & Court, D. *A. Rev. Genet.* **13,** 319–353 (1979).
38. Piperno, J. R. & Oxender, D. L. *J. biol. Chem.* **243,** 5914–5920 (1968).
39. Gilbert, W. & Maxam, A. *Proc. natn. Acad. Sci. U.S.A.* **70,** 3581–3584 (1973).
40. Steitz, T. A., Richmond, T. J., Wise, D. & Engelman, D. *Proc. natn. Acad. Sci. U.S.A.* **71,** 593–597 (1974).
41. McGinnis, E. & Williams, L. *J. Bact.* **111,** 739–744 (1972).
42. McGinnis, E. & Williams, L. *J. Bact.* **109,** 505–511 (1972).
43. Pizer, L. I., McKitrick, J. & Tosa, T. *Biochem. biophys. Res. Commun.* **49,** 1351–1357 (1972).
44. Clarke, S. J., Low, B. & Konigsberg, W. *J. Bact.* **113,** 1096–1103 (1973).
45. Baer, M., Low, K. B. & Söll, D. *J. Bact.* **139,** 165–175 (1979).
46. La Rossa, R., Vogeli, G., Low, K. B. & Söll, D. *J. molec. Biol.* **117,** 1033–1048 (1977).
47. Hirshfield, I. N., Liu, C. & Yeh, F. *J. Bact.* **131,** 589–597 (1977).
48. Wirth, R., Buckel, P. & Böck, A. *FEBS Lett.* **83,** 103–106 (1977).
49. Cashel, M. & Gallant, J. in *Ribosomes* (eds Nomura, M., Tissieres, A. & Lengyel, P.) 733–745 (Cold Spring Harbor Laboratory, New York, 1974).
50. Block, R. & Haseltine, W. A. in Ribosomes (eds Nomura, M., Tissieres, A. & Lengyel, P.) 747–761 (Cold Spring Harbor Laboratory, New York, 1974).
51. Morse, D. E. & Morse, A. N. *J. molec. Biol.* **103,** 209–226 (1976).
52. Stephens, J. C., Artz, S. W. & Ames B. N. *Proc. natn. Acad. Sci. U.S.A.* **72,** 4389–4393 (1975).
53. Yates, J. L., Arfsten, A. E. & Nomura, M. *Proc. natn. Acad. Sci. U.S.A.* **77,** 1837–1841 (1980).
54. Fiil, N. P., Friesen, J. D., Downing, W. L. & Dennis, P. P. *Cell* **19,** 837–844 (1980).
55. Dean, D. & Nomura, M. *Proc. natn. Acad. Sci. U.S.A.* **77,** 3590–3594 (1980).
56. Yates, J. L. & Nomura, M. *Cell* **21,** 517–522 (1980).
57. Zengel, J. M., Mueckl, D. & Lindahl, L. *Cell* **21,** 523–535 (1980).
58. Maxam, A. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74,** 560–564 (1977).
59. Berkner, K. L. & Folk, W. R. *J. biol. Chem.* **252,** 3176–3189 (1977).

LETTERS TO NATURE

Newtonian gravity measurements impose constraints on unification theories

G. W. Gibbons & B. F. Whiting

Department of Applied Mathematics and Theoretical Physics,
University of Cambridge, Silver Street, Cambridge CB3 9EW, UK

Theories which attempt to unify gravity with the other forces of nature can be coarsely classified according to the mass scale of the new particles they introduce or equivalently the length scale at which new phenomena occur. This mass scale can be expressed as $m_p(m_H/m_p)^n$, where m_H is a typical hadron mass and m_p is the Planck mass. In most current theories n is 0 or 1. However, in some theories $n=2$, which offers the possibility of experimental consequences at kilometre scales. Here using satellite and geophysical data we place constraints on such theories and find that they are not viable unless $m_H > 10^3$ GeV.

These theories[1-7] predict deviations from Newton's inverse square law at length scales $L \simeq (m_p/m_H)\hbar/m_H c$ where m_p is the Planck mass

$$((c\hbar/G)^{1/2} \simeq 2.2\times 10^{-5}\,\mathrm{g} = 1.2\times 10^{19}\,\mathrm{GeV})$$

An L of $\simeq(1\,\mathrm{GeV}/m_H)^2$ km is the Compton wavelength of a particle of mass

$$\mu = (m_H/1\,\mathrm{GeV})^2\,10^{-10}\,\mathrm{eV} \simeq (m_H/1\,\mathrm{GeV})^2\,10^{-6}\,\mathrm{K}$$

If m_H is the neutron mass we obtain the order of magnitude for the radius of a neutron star. In the theories of Fujii[1-4] where the masses of elementary particles and the dimensional nature of the gravitational coupling constant G have their common origin in a spontaneous breakdown of dilation invariance, μ is the mass of a scalar particle which couples to ordinary matter with gravitational strength. Zee[5,6] derives Newton's constant in a similar way, but the mass of the resultant scalar particle in his theory would seem to be much heavier than the masses envisaged here. Scalars with low mass can also be a consequence of the super-Higgs effect in supergravity theories. In extended supergravity theories the further possibility arises of vector particles with a mass $\simeq\mu$. These would also couple to ordinary matter with gravitational strength but would give rise to a repulsive force[9-12]. Note that Majorana Fermion mass terms of this magnitude have been suggested for neutrinos arising from a gravitationally-induced four-Fermi interaction[13] and for the spin 3/2 gravitino particle in supergravity theories through a supersymmetric Higgs Mechanism[14]. Because of supersymmetry one might expect bosonic partners with mass of order μ. Such a spin 1 partner for the gravitino has been suggested by Fayet[15].

Length scales of 1 km lie between the distance range of laboratory measurements of Newton's constant, G, and astronomical distances over which Newton's law has most accurately been checked. It has been argued[16] that our experimental knowledge of gravitational forces between 1 m and 10 km is so poor that it allows a considerable difference between the laboratory measured gravitational constant and its value on astronomical scales—an effect predicted in theories of the type alluded to above. Even at laboratory scales Long[17] has reported non-newtonian behaviour. The null experiment of Spero *et al.*[18] conflicts with Long's result unless some mechanism like that suggested by Long[19] operates. However, conventional vacuum polarization effects in quantum gravity do not lead to the behaviour required by Long[19]: such effects are insignificant[20]. We point out that existing data from satellite and geodesy measurements[21,22] together with the results of a recent mine experiment[23] and those in refs 18, 24 place stringent limits on the parameters appearing in theories of the Fujii or Scherk types.

If gravitation is due to the exchange of particles of Compton wavelength λ_i, and coupling α_i, the potential energy of two masses m and m' at a separation r is V where

$$V = -G_\infty \frac{mm'}{r}\left(1+\sum_{i=1}^{N} \alpha_i \exp(-r/\lambda_i)\right) \tag{1}$$

As the additional Yukawa terms are negligible for $r \gg \max \lambda_i$, G_∞ is the gravitational constant at very large distances. For $r \ll \min \lambda_i$ the newtonian law also applies with a gravitational constant

$$G_0 = G_\infty\left(1+\sum_{i=1}^{i=N} \alpha_i\right)$$

Negative values of α_i correspond to repulsions caused by the exchange of vector particles and positive α_i to attractions due to the exchange of scalar particles. Potentials of the form of equation (1) occur in several scalar–tensor theories of gravity[25,26] in which the mass of the scalar particle remains unspecified. Such potentials also arise in higher derivative theories of gravity[27-29], the mass usually being thought of as having a planckian or possibly hadronic value. According to De Witt[28] the value is to be determined experimentally. Stelle[29] has shown that theories predict $\alpha = \frac{1}{3}$ for the massive scalar and $\alpha = -\frac{4}{3}$ for the massive tensor ghost particle. Fujii's first theory[1] also predicted $\alpha = \frac{1}{3}$.

For the present purposes we consider $N = 1$, so that

$$V = -G_\infty \frac{mm'}{r}(1+\alpha \exp(-r/\lambda)) \tag{2}$$

The effective gravitational constant $G(r)$ is given by

$$G(r) = G_\infty(1+\alpha(1+r/\lambda)\exp(-r/\lambda)) \tag{3}$$

A measured value of the ratio $G(r_1)/G(r_2) = \beta$ at the distances r_1 and r_2 constrains α and λ to lie on a curve in (α, λ) space given by

$$\alpha = \frac{(\beta-1)}{[(1+r_2/\lambda)\exp(-r_2/\lambda) - \beta(1+r_1/\lambda)\exp(-r_1/\lambda)]} \tag{4}$$

Varying β within the limits of experimental errors causes the curve to sweep out an allowed region in the (α, λ) plane. Spero *et al.*[18] used this procedure to obtain the region allowed by their results and by those of Long[17]; these are plotted in Fig. 1 for positive and negative α. Figure 1 also shows the regions allowed by the experiment of Panov and Frontov[24], which effectively restrict $|\alpha|$ to be $<10^{-2}$ out to 3 m. (The results of Yu *et al.*[30] and Hirakawa *et al.*[31] provide poorer limits and are not included.)

These experiments constrain $G(r)$ at laboratory scales. Mikkelsen and Newman have provided limits from lunar surface gravity and Mercury and Venus flyby data which constrain $G(r)$ at the upper end of the range we are considering. These are also shown in Fig. 1 from which it can be seen that for $3m < \lambda < 10^3$ km α is very poorly constrained. To fill this gap we turn to the data available from satellites and geodesy[21,22], using essentially independent computations of equatorial gravity from an extensive net of gravity anomaly measurements and from satellite data for the Earth's potential. Mikkelsen and Newman regarded data of this sort as unreliable; they instead considered lunar surface gravity data which, because of the smaller size of the Moon, gives a much tighter constraint on α for a given error of measurement. We believe that more recent data avoids the objections raised for the Earth in ref. 16.

The potential at a distance r from the centre of a uniform sphere of density ρ and radius R due to the Yukawa contribu-

0028-0836/81/260636—03$01.00

tion to gravity is

$$\frac{4\pi}{3}\cdot\frac{G_\infty \rho R^3}{r}\alpha f(R/\lambda)\exp\{(R-r)/\lambda\} \qquad (5)$$

where

$$f(x)=3e^{-x}(x\cosh(x)-\sinh(x))x^{-3} \qquad (6)$$

is a 'form factor' which takes into account the fact that only a region of the order of λ from the field point contributes to the Yukawa part of gravity. For a non-uniform sphere ρ should be replaced by a suitably averaged density $\bar{\rho}$ within a few λ of the field point. The acceleration due to gravity at a height h above the surface of the Earth is therefore

$$g(h)=G_\infty M_\oplus(R+h)^{-2}$$
$$\times[1+y\alpha f(R/\lambda)(1+(R+h)/\lambda)\exp(-h/\lambda)] \qquad (7)$$

where $M_\oplus=4\pi\rho R^3/3$ is the mass of the Earth (ρ now being the mean density) and $y=\bar{\rho}/\rho$. On the surface ($h=0$)

$$g(h)\rightarrow g(0)=G_\infty MR^{-2}[1+y\alpha f(R/\lambda)(1+R/\lambda)] \qquad (8)$$

At a height $h\gg\lambda$ we have

$$g(h)\rightarrow g_\infty(r)=G_\infty M(R+h)^{-2} \qquad (9)$$

Thus the extrapolation from measurements of gravity at large heights above the Earth to the surface will produce a fractional error

$$\varepsilon=(g_{\text{measured}}-g_{\text{extrapolated}})/g_{\text{extrapolated}} \qquad (10)$$
$$=\alpha y f(R/\lambda)(1+R/\lambda) \qquad (11)$$

If $R\gg\lambda$ this gives

$$\varepsilon=3\alpha\lambda y/2R \qquad (12)$$

The accuracy with which terrestial and satellite measurements of gravity agree has been discussed by Rapp[21,22] who provides estimates of normal gravity at the Equator, γ_e, derived from satellite models, ranging from 978,031.3 to 978,033.0 mGal (1 Gal = 1 cm s^{-2}) and also determines γ_e, from an extensive net of 1°×1 mean free air anomalies, as 978,030.9 mGal with an estimated standard error of ±1 mGal. Thus satellite and terrestial values agree to within ~2 p.p.m. Rapp's adjusted value of γ_e, 978,031.69 mGal is based on a best set of satellite data and includes his geodetic determination of γ_e. His computation (R. H. Rapp, personal communication) enables the dependence of this adjusted value on the geodetic data to be removed and the resultant estimate, $\gamma_e=978{,}032.085$ mGal agrees with the geodetic data to within 1.2 mGal. All Rapp's calculations have removed the effect of the atmosphere. Thus:

$$\varepsilon=(978{,}030.9-978{,}032.085)/978{,}032.085=-1.2\times10^{-6}$$

Rapp has suggested (personal communication) that current estimates of the agreement may be a factor of 10 better than this.

Mikkelsen and Newman's[16] objected to geophysical data because the local agreement between measured and calculated gravity anomalies may be a factor of 10^3 worse than the agreement of these globally averaged quantities. This is too pessimistic; Rapp[22] shows that the agreement is within 10 mGal.

Rapp computes the potential coefficients in a multipole expansion of the Earth's gravitational field from 5°×5° mean gravity anomalies and compares his results with the coefficients of the Earth model GEM 7 which were obtained using satellites. The mean square difference in gravity anomalies, δg, calculated allowing for various terrain corrections, depended on the number of harmonics used. For a maximum degree of 3, δg did not exceed 2.2 mGal; as the maximum degree increased these differences increased, but even for a maximum degree of 16 δg never exceeded 7.3 mGal. These relatively small discrepancies for local anomalies indicate that an agreement of as little as 10^{-6} is reasonable for a globally averaged quantity and a discrepancy larger than 10^{-5} is unacceptable. The constraints obtained in both these cases are shown in Fig. 1. The larger value of the Earth's radius is more than compensated for by the greater precision of the measurements compared with the lunar surface gravity measurements. The constraint, $\alpha<10^{-2}$, applies if $\lambda>$ 1 km for $\varepsilon=10^{-6}$, and if $\lambda>10$ km for $\varepsilon=10^{-5}$.

We now discuss the effect of the Yukawa addition to gravity on the variation of gravity with depth below the surface of the Earth. This variation may be measured by experiments in mines[23] or in submarines[32]. The acceleration due to gravity at a depth d below the surface of a homogeneous sphere of radius R and density ρ is

$$g(R-d)=\alpha\frac{4\pi}{3}G_\infty\rho(R-d)\times$$
$$\times(1+R/\lambda)\exp(-d/\lambda)f((R-d)/\lambda) \qquad (13)$$

Assuming λ/R and d/R are both small we obtain

$$g(R-d)-g(R)=2\pi G_\infty\rho\alpha\lambda(\exp(-d/\lambda)-1) \qquad (14)$$

When applying equation (14) to an inhomogeneous Earth the appropriate density ρ is ρ_m, obtained as some average of the density of the material through which the descent is made, including material to a distance λ below the depth d.

The standard newtonian change in gravity is given by

$$g_N(R-d)-g_N(R)=-\frac{4\pi}{3}G_\infty(3\rho_M-2\rho_\oplus)\,d \qquad (15)$$

where $\rho_\oplus$ is the mean density of the Earth.

Equation (15) is accurate to an order of (d/R). To this order of accuracy equation (5) or (13) both give negligible contribution and give, for gravity at the surface of the Earth,

$$g_0=g(R)=\frac{4\pi}{3}G_\infty\rho_0R \qquad (16)$$

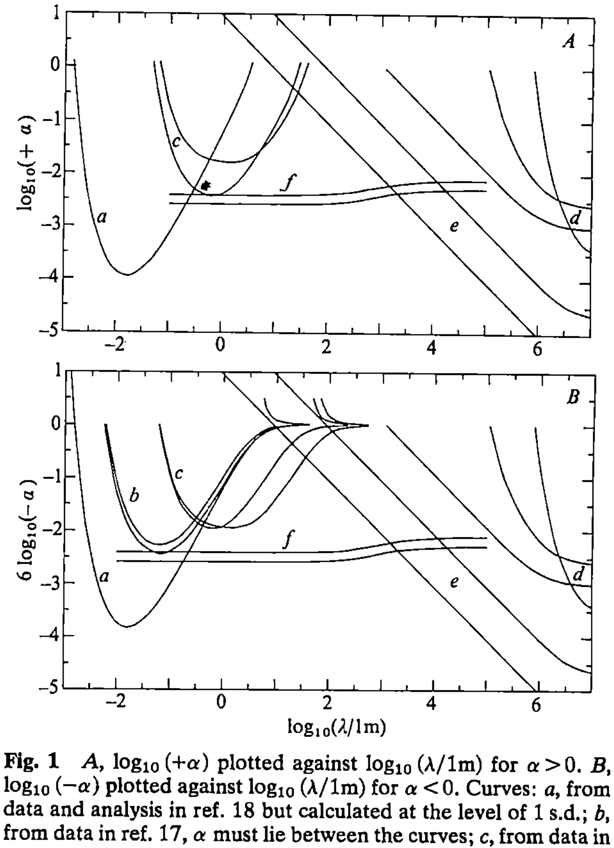


Fig. 1 *A*, $\log_{10}(+\alpha)$ plotted against $\log_{10}(\lambda/1\text{m})$ for $\alpha>0$. *B*, $\log_{10}(-\alpha)$ plotted against $\log_{10}(\lambda/1\text{m})$ for $\alpha<0$. Curves: *a*, from data and analysis in ref. 18 but calculated at the level of 1 s.d.; *b*, from data in ref. 17, α must lie between the curves; *c*, from data in ref. 24 for two separate ranges of experiment; *d*, taken from ref. 16, giving curves obtained from results for lunar surface gravity, and Mercury and Venus fly-bys; *e*, resulting from comparison of satellite and geodesy data[21]; the upper curve is assuming there is agreement to 0.1 p.p.m., the lower 1 p.p.m.; *f*, calculated for a mine experiment to 1% accuracy (upper curve) and a submarine experiment to 0.1% accuracy, both to a depth of 1 km. Neither curve can be extended indefinitely for small or large λ. Note that α is restricted to lie below each curve.

Expressing results in terms of laboratory measured values of Newton's constant, $G_0 = (1+\alpha)G_\infty$, and surface gravity, g_0, the total change in gravity is

$$\Delta g(\alpha, d) = -4\pi d\left(\frac{\rho_M G_0}{1+\alpha} - \frac{g_0}{2\pi R}\right) - \frac{\alpha}{1+\alpha} 2\pi G_0 \rho_M \lambda (1 - \exp(-d/\lambda)) \qquad (17)$$

If ε is the fractional difference between this and the newtonian value,

$$\varepsilon = (\Delta g(\alpha, d) - \Delta g(0, d))/\Delta g(0, d) \qquad (18)$$

An experiment giving a value for ε again defines a curve in the (α, λ) plane according to

$$\alpha = -\varepsilon\left(1 - \frac{g_0}{2\pi G_0 \rho_M R}\right) \Big/ \left(\varepsilon\left(1 - \frac{g_\infty}{2\pi G_\infty \rho_M R}\right) + 1 - \frac{\lambda}{2d}(1 - \exp(-d/\lambda))\right) \qquad (19)$$

Experimental limits on ε then confine the theory to a restricted region of the (α, λ) plane. Figure 1 shows the curves obtained by setting $\varepsilon = 0.01$, $\rho = 2.65$ as a representational value for crustal rock and $\varepsilon = 0.001$, $\rho = 1.03$ for seawater, assuming an experiment carried out to a depth of 1 km. Mine experiments accurate to only a few per cent can give very useful limits to α in the range of λ where α is otherwise poorly constrained. Because of factors related to the density differences involved a submarine experiment to the same depth as a mine experiment will need ~10 times greater accuracy before it can put a better constraint on α. Stacey has reported results from an experiment to a depth of 950 m in a mine at Mount Isa, Queensland giving $|\varepsilon| < 0.05$.

We conclude that there is very little scope for a theory which allows deviations of >1% from Newton's law of gravitational attraction on laboratory or larger length scales. Good mine or submarine experiments will further constrain these theories. Further large scale experiments are essential to improve bounds on α between 1 m and 10 km.

We thank A. H. Cook, D. Gubbins, S. W. Hawking, and R. Newman for useful discussions and F. Stacey, P. Bender and R. H. Rapp for useful correspondence.

Received 3 November 1980; accepted 24 April 1981.

1. Fujii, Y. *Nature phys. Sci.* **234,** 5–7 (1971).
2. Fujii, Y. *Ann. Phys.* **69,** 494–521 (1972).
3. Fujii, Y. *Phys. Rev.* D9, 874–876 (1974).
4. Fujii, Y. *Gen. Relat. Gravitat.* **6,** 29–34 (1975).
5. Zee, A. *Phys. Rev. Lett.* **42,** 417–521 (1979).
6. Zee, A. *Phys. Rev. Lett.* **44,** 703–706 (1979).
7. Ellis, J., Galliard, M. K. & Zumino, B. *Phys. Lett.* **94B,** 343–348 (1980).
8. Fujii, Y. & Nishino, N. *Z. Phys. C.* **2,** 247–252 (1979).
9. Scherk, J. *Phys. Lett.* **88B,** 265–267 (1979).
10. Scherk, J. in *Supergravity* (eds Freedman, D. & van Nieuwenhuizen, P.) 43–51 (North-Holland, Amsterdam, 1979).
11. Scherk, J. in *Mathematical Problems in Theoretical Physics* (ed. Osterwalder, K.) (Springer Lecture Notes in Physics No. 116, 1980).
12. Scherk, J. in *Unification of the Fundamental Particle Interactions* (eds Ferrara, S., Ellis, J. & van Nieuwenhuizen, P.) 381–410 (Plenum, New York, 1981).
13. Barber, R., Ellis, J. & Galliard, M. K. *Phys. Lett.* **90B,** 253–257 (1980).
14. Deser, S. & Zumino, B. *Phys. Rev. Lett.* **38,** 1433–1436 (1977).
15. Fayet, P. *Phys. Lett.* **95B,** 285–289 (1980).
16. Mikkelsen, D. R. & Newman, N. J. *Phys. Rev.* **D16,** 919–926 (1977).
17. Long, D. R. *Nature* **260,** 417–418 (1976).
18. Spero, R. *et al. Phys. Rev. Lett.* **44,** 1645–1648 (1980).
19. Long, D. R. *Nuovo Cim.* **85B,** 252–256 (1980).
20. Duff, M. J. *Phys. Rev.* **D9,** 1837–1839 (1974).
21. Rapp, R. H. *Geophys. Res. Lett.* **1,** 35–38 (1974).
22. Rapp, R. H. *Bull. Geod.* **51,** 301–323 (1977).
23. Stacey, F. D. *et al. Phys. Rev.* **D23,** 1683–1692 (1981).
24. Panov, V. I. & Frontov, V. N. *Soviet Phys. JETP* **50,** 852–856 (1979).
25. O'Hanlon, J. *Phys. Rev. Lett.* **29,** 137–138 (1972).
26. Acharya, R. & Hogan, P. A. *Lett. Nuovo Cimento* **6,** 668–672 (1973).
27. Pechlana, E. & Sexl, R. *Commun. Math. Phys.* **2,** 165–175 (1966).
28. de Witt, B. *Phys. Rep.* **19C,** 295–357 (1975).
29. Stelle, K. *Gen. Relat. Gravitat.* **9,** 353–371 (1978).
30. Yu, H. T. *et al. Phys. Rev.* **D20,** 1813–1815 (1979).
31. Hirakawa, H. *et al. Nature* **283,** 184–185 (1980).
32. Stacey, F. *Geophys. Res. Lett.* **5,** 377–378 (1978).

0028–0836/81/260638—03$01.00

Meteoric compounds in the stratosphere

A. C. Aikin

Laboratory for Planetary Atmospheres, Goddard Space Flight Center, Greenbelt, Maryland 20771, USA

The presence of gaseous metal oxides and hydroxides originating from incoming meteoric material has been suggested to explain stratospheric ion composition[1] and as chemical sinks for chlorine species[2]. Recent ion composition data can be explained without recourse to metal hydroxides[3–5], although the results do not preclude gaseous metal compound concentrations of 10^4 cm^{-3} at 35 km. However, theoretical considerations suggest that gaseous meteoric material coagulates to form particulate matter[6]. No direct measurements of metal compounds exist. Photoionization of meteoric compounds such as NaOH and CaOH in the 40–70-km altitude region produces a diurnally varying source of ionization in addition to nearly constant cosmic ray ionization. Ion pair production can be several times that due to cosmic radiation depending on sunspot cycle, altitude, time of day and meteor activity. It is suggested here that the behaviour of the ion density should be studied to determine the existence of gaseous metal oxides. The available ion density and conductivity data are insufficient to demonstrate conclusively the existence of such compounds, although unexplained increases are observed. The presence of meteoric compounds, would imply an electrical conductivity which is related to both solar and meteor activity and would thus be an important factor in the physical processes involved in those Sun–weather relationships which depend on atmospheric electrical conductivity.

Table 1 lists several of the more abundant elements contained in meteoric material. The abundance is given relative to silicon, which is taken to be 100. Also listed are compounds, which are likely to be formed as the meteoric element passes through the atmosphere. The wavelength equivalent of the ionization threshold of the compounds is also given. Except for NaOH these values are based on the compilation by Rosenstock *et al.*[7]. The ionization potential for NaOH is estimated by adding H to NaO for which the ionization threshold is known: this addition should not raise the ionization threshold. Candidates for ionization below 70 km include CaO, CaOH, NaO and NaOH. In the case of Na the hydroxyl, NaOH, is predicted to be the principal compound below 80 km chiefly because of the reactions

$$\mathrm{Na} + \mathrm{O_2} + \mathrm{M} \rightarrow \mathrm{NaO_2} + \mathrm{M}$$
$$\mathrm{NaO_2} + \mathrm{OH} \rightarrow \mathrm{NaOH} + \mathrm{O_2}$$

Table 1 Elemental abundance and ionization threshold of meteoric compounds

Element	Abundance	Compound	Ionization threshold (Å)
Si	100	SiO	1,062.5
Fe	74	FeO	1,550.0
Ni	3.9	NiO	1,305.3
Cr	1.1	CrO	1,476.2
Al	5.9	AlO	1,305.3
Na	4.3	NaO	1,907.7
		NaOH	1,907.7
K	0.3	KOH	1,653.3
Ca	5.5	CaO	1,907.7
		CaOH	2,101.7

All values except for NaOH are based on the compilation in ref. 7.

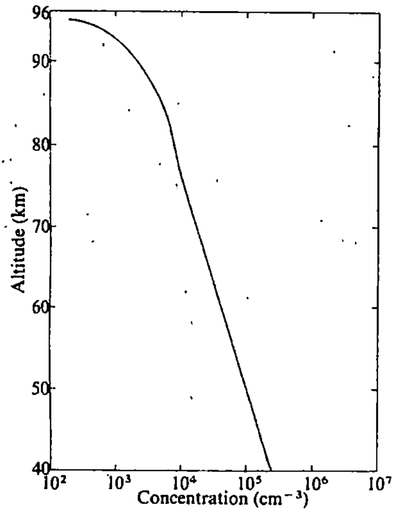


Fig. 1 Adopted altitude distribution of NaOH or CaOH.

Calculations for atomic sodium diffusing downward from its observed peak at 90 km, while reacting photochemically but with no loss due to particulate formation, are in general agreement with similar calculations[8]: NaOH is the principal sodium compound.

Because calcium can have nearly the same meteoric abundance as sodium[9], the total density distribution of calcium compounds will follow that of sodium (neutral chemistry for calcium has not been established). If a reaction sequence analogous to that for sodium occurs, then one expects CaOH to be the principal component below 70 km. For calculations of the ionization rate the distribution shown in Fig. 1 will be adopted and be assumed to apply to either CaOH or NaOH.

The photoionization production rate for a constituent of concentration $[M]$ is derived using the expression

$$q = 10^{-18} \sum_{\lambda > \lambda_0} [M] Q_\lambda e^{-\tau_\lambda}$$

where Q_λ is the solar flux at wavelengths less than the ionization threshold, λ_0 (1,908 Å for NaOH, 2,100 Å for CaOH). The optical depth τ_λ is the product of column densities and absorption cross-sections. Ozone and molecular oxygen are the absorbers. At 2,000 Å there is minimum in the O_3 cross-section which allows the penetration of this radiation to photoionize CaOH. The present calculation is for a solar zenith angle of 30°. Photoionization cross-section is assumed to be 10^{-18} cm^2. The total ion-pair production function shown in Fig. 2 has diurnal variation superimposed on the nearly constant cosmic-ray ionization rate. Ionization by cosmic radiation varies by a factor of two with solar cycle and an order of magnitude with geomagnetic latitude[10,11]. Both the minimum amount of cosmic-ray ionization, geomagnetic equator at sunspot maximum, and nearly the maximum amount, high latitude sunspot minimum, are given in Fig. 2. Any photoionization of meteoric natural adds to the cosmic radiation. Solar Lyα ionization of nitric oxide is the ionizing agent above 70 km.

For circumstances at low geomagnetic latitudes $CaOH^+$ and $NaOH^+$ can be dominant ions formed first by ionization. At high geomagnetic latitudes the influence of photoionization will be minimized due to the combination of lower average solar zenith angles and larger cosmic-ray ionization.

The hypothesis that meteoric material is adding to the ionization level of the atmosphere below 70 km is consistent with observations. Widdel *et al.*[12] have observed an increase in ion density above 40 km with a maximum of 10^4 cm^{-3} at 50 km. The loss rate in this altitude region is due to ion–ion recombination between positive and negative ions. Smith *et al.*[13] have measured the ion–ion recombination rate for negative ions observed in the atmosphere and water cluster ions. A mean value of 5×10^{-8} cm^3 s^{-1} is observed. No measurements are available for ion–ion recombination involving meteor material-based ions, but a value in the 5×10^{-8} cm^3 s^{-1} range cannot be excluded. The observed ion density cannot be explained by cosmic radiation alone using this value of the ion–ion recombination coefficient. The ion pair production rate will also vary in direct proportion to the meteor flux.

Ion composition measurements show that water cluster ions predominate to 55 km with other ions, increasing in importance at lower altitudes so that by 40 km 50% or more of the ions are not $H^+ \cdot (H_2O)_n$. If $CaOH^+$ or $NaOH^+$ are formed in large quantities then loss mechanisms must transfer these ions to match the observed composition. Possible reactions include

$$NaOH^+ + H_2O \rightarrow H_3O^+ + NaO$$

$$CaOH^+ + H_2O \rightarrow H_3O^+ + CaO$$

The ion composition measurements above 50 km were conducted at a geomagnetic latitude of 69° often at high solar zenith angles. These are precisely the conditions where the contribution by photoionized meteoric material is expected to be minimal so that more H_3O^+ is expected.

An additional constraint is that the densities of ionizable constituents such as NaO, CaO, CaOH and NaOH must be low enough in the *D* region so that Lyα ionization remains the primary mechanism for ion formation between 70 and 85 km. Diurnal studies of electron density support Lyα. Sunrise variations are more difficult to interpret due to photodetachment of negative ions and lack of precise data. Long-wave UV radiation will penetrate to 70 km before Lyα and produce ionization at an earlier time. The data of Mitchell *et al.*[14] indicate anomalous increases in conductivity at sunrise. Data of Croskey *et al.*[15] show an increase in equatorial conductivity above 60 km compared with mid-latitude values. Both these measurements are consistent with an additional source of ionization.

There is, therefore, no direct measurement of NaOH or other metal compounds postulated by Murad *et al.*[2] to react chemically with stratospheric chlorine. Anomalous increases observed in ion densities and the sunrise behaviour of conductivity may be an indication of the existence in the mesosphere and stratosphere of such compounds, as they can be photoionized by UV radiation which penetrates to these heights. A detailed study of the ion density variations particularly at low latitudes near sunspot maximum is required to establish the existence of metal hydroxides. The presence of such compounds will have consequences for Sun–weather relationship theories based on electrical conductivity. It has been argued that variations in electrical conductivity affect the electric field between clouds and the Earth and lead to changes in cloud electrification, which

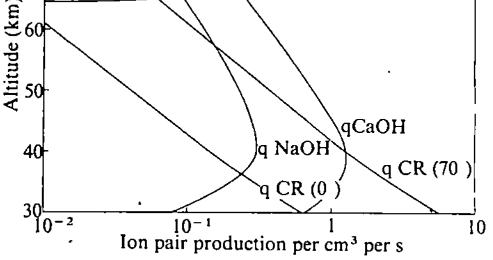


Fig. 2 Ion-pair production rate for the middle atmosphere. Solar zenith angle is 30°. A maximum and minimum cosmic-ray (CR) ionization rate is shown. Top curve is photoionization of NO.

can affect thunderstorm activity and weather[16]. If photoionization of meteoric material controls conductivity between 40 and 70 km then solar activity variations of conductivity would be expected. More importantly the ionization rate will be affected not only by solar radiation variability but also by meteor flux.

We thank J. Herman and M. Krauss for useful discussions.

Received 5 March; accepted 29 April 1981.

1. Ferguson, E. C. *Geophys. Res. Lett.* **5**, 1035–1038 (1978).
2. Murad, E., Swider, W. & Benson, S. W. *Nature* **289**, 273–274 (1981).
3. Arijs, E., Nevehans, D. & Ingels, J. *Nature* **288**, 684–686 (1980).
4. Arnold, F. & Henschen, G. *Geophys. Res. Lett.* **8**, 83–86 (1981).
5. Arnold, F., Henschen, G. & Ferguson, E. E. *Planet. Space Sci.* **29**, 185–193 (1981).
6. Hunten, D. M., Turco, R. P. & Toon, O. B. *J. atmos. Sci.* **37**, 1342–1357 (1980).
7. Rosenstock, H. M., Draxl, K., Steiner, B. W. & Herron, J. T. *J. phys. Chem. Ref. Data* **6**, Suppl. No. 1, (1977).
8. Liu, S. C. & Reid, G. C. *Geophys. Res. Lett.* **6**, 283–286 (1979).
9. Goldberg, R. A. & Aikin, A. C. *Science* **180**, 294–295 (1973).
10. Nicolet, M. & Aikin, A. C. *J. geophys. Res.* **65**, 1469–1485 (1960).
11. Heaps, M. G. *Planet. Space Sci.* **26**, 513–517 (1978).
12. Widdel, H. U., Rose, G. & Borchers, R. *J. geophys. Res.* **81**, 6217–6220 (1976).
13. Smith, D., Adams, N. G. & Church, M. J. *Planet. Space Sci.* **24**, 697–703 (1976).
14. Mitchell, J. D., Sagar, R. S. & Olsen, R. O. *Space Research* Vol. 17 (ed. Rycroft, M. I.) (Pergamon, Oxford, 1976).
15. Croskey, C. L., Hale, L. C. & Leider, S. C. *Space Research* Vol. 17 (ed. Rycroft, M. I.) (Pergamon, Oxford, 1976).
16. Markson, R. & Muir, M. *Science* **208**, 979–990 (1980).

A physicochemical mechanism for the ignition of the Seveso accident

T. G. Theofanous

School of Nuclear Engineering, Purdue University, West Lafayette, Indiana 47907, USA

The so-called Seveso 'explosion' of 10 July 1976 was completely unexpected, and various speculations about its initiation have been entertained, including the possibility of sabotage[1]. Seveso has remained the focus of undiminished interest and publicity[1,2] and the puzzle concerning the occurrence of its accident persists. The explanation proposed here is based on our recent investigations of the relevent chemical kinetics and heat transfer phenomena. This physicochemical mechanism involves the combination of (1) exothermic reactions starting (personal communication from the Givaudan Company) below the previously known 230 °C threshold and (2) a thermal conduction–convection–radiation process responsible for channelling the normally negligible (in global sense) upper reactor wall heat into a small portion (·10%) of the reactor contents, thus yielding the temperature range conducive to these exothermic reactions.

At the time of the accident the Seveso reactor contained 2,030 kg sodium trichlorophenate, 540 kg sodium chloride, ~1,000 kg ethylene glycol, and 1,600 kg sodium ethylene glycolate, diethylene glycol and polyethylene glycol. Such mixtures are known[3] thermally to decompose above 230 °C by virtue of an exothermic chemical reaction yielding gaseous byproducts as well as the highly toxic 2,3,7,8-tetrachlorodibenzodioxin (TCDD). Because such accidents have occurred several times previously, special precautions were taken to prevent the Seveso reactor from reaching this dangerous temperature range. The pressure of the steam utilized for heating (to distill the ethylene glycol) was limited to 12 bar and thus a saturation temperature of 190 °C. Heating was interrupted, and air admitted to re-establish atmospheric pressure, stirring of the reactor contents continued for 15 min and the reactor was left with its liquid contents at the safe temperature level of 158 °C. The accident occurred 7.5 h later.

Early investigations[1] were directed at evaluating the long-term behaviour of reactor-like solutions at various temperatures around the operating level of 158 °C. In particular the potential for slow, slightly exothermic reactions promoted by air (oxidation), or various metal surfaces (catalysis) was thoroughly explored. The results of these studies (indicating stable behaviour) together with recent developments leading to an expectation of local heating (of a small portion of the reactor contents) up to 200–220 °C stimulated interest in a more careful investigation of this temperature range. Differential thermal analysis (DTA) of typical reactor content samples[4,5] identified two 'slow' exothermic reactions occurring before the decomposition reaction. The first initiates at ~185 °C, peaks at 235 °C and is characterized by an adiabatic temperature rise of 57 °C. The second exotherm initiates at ~255 °C, overlapping with the tail end of the first one, peaks at 265 °C and its adiabatic temperature rise is estimated at 114 °C. The activation energies and the logarithms of the frequency factors are 24,000 K, 29,800 K and 40 and 48 respectively. The decomposition reaction is initiated at ~280–290 °C and exhibits a rapid pressure buildup at ~300 °C. Its adiabatic temperature is estimated at 142 °C.

The adiabatic induction times are 2.1 h and 30 min for the first (~200 °C) and second (~240 °C) exotherm respectively. It is easy to see how the decomposition reaction may be reached quite rapidly (compared with the Seveso time scale) in an adiabatic fluid volume that has been heated to ~200 °C. However, the actual thermal process driving the ignition is approximated well by the air–liquid interface held at a constant temperature in the 200–220 °C range. In such conditions the temperature gradients may be steep, the heat losses may be substantial and the thermal ignition theory[6] would predict times for self-accelerating conditions substantially longer than the adiabatic induction times (roughly 10 times longer). However, in relation to this theory the actual situation is complicated by (1) three consecutive, and partly overlapping reactions (the first two, for which reactant depletion cannot be neglected) serve for warmup; and (2) uncommon and complication-yielding boundary conditions during the warmup period (that is, assuming the surface temperature to a constant value during this period accentuates heat losses and produces overestimation of ignition times). We must regard, therefore, the thermal ignition theory results only as qualitative estimates indicating that: (1) heatup of the liquid surface to ~200–220 °C represents a credible initiation of the decomposition-yielding chain of reactions within ~7 h; and (2) heat losses and reactant depletion need be taken into account for a more precise evaluation of the warmup

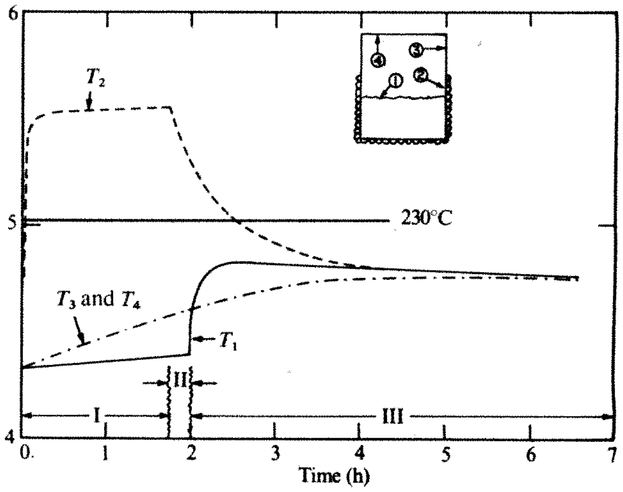


Fig. 1 Calculated thermal transient of the Seveso reactor walls and reaction mixture free surface. Inset is schematic of the Seveso reactor. Surface 2 corresponds to the unsubmerged heated surface (backed by steam coils). Emissivities of 1 and 0.5 were used for the liquid and stainless wall respectively.

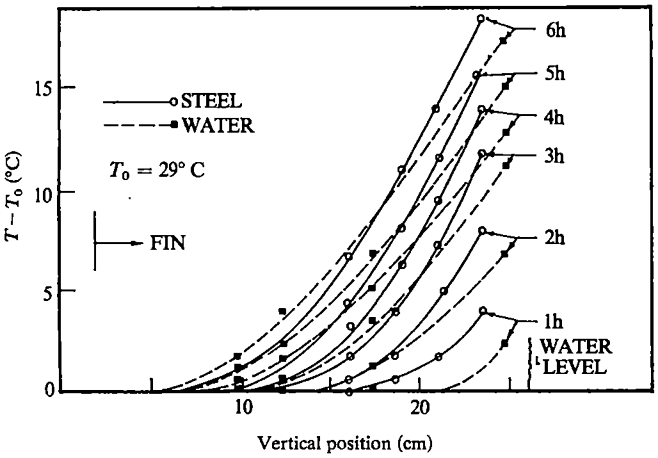


Fig. 2 Experimental demonstration of the stratification phenomena in the submerged fin case with a driving temperature gradient of ~10 °C cm^{-1}. Lateral temperature gradients in the liquid were negligible.

period. The thermal-physical processes driving the initial stages of the ignition are now considered.

The geometry of the reactor, the heating jacket and the liquid level are indicated in Fig. 1 (inset). Distillation of ethylene glycol was carried out under vacuum (10 torr) by means of 12 bar superheated (~300 °C) steam. At the time that heating was interrupted the distillation had proceeded for 1$\frac{3}{4}$ h (period I). During this period surface 2 (I_2) heats quickly to a temperature near that of the steam. The liquid remains at a uniform temperature of ~160 °C and surfaces 3 and 4 (I_3 and I_4) experience slow heating due to thermal radiation from surface 2. During the subsequent 15 min (period II) with continuing agitation the liquid remains at the uniform temperature of ~160 °C, the submerged reactor wall cools also to this same temperature, surfaces 3 and 4 continue to heat up slowly while surface 2 rapidly cools by radiation to all three other surfaces. Clearly in the neighbourhood of surface junctions steep temperature gradients and additional cooling will exist; however, a reasonable estimate of temperature transients may be obtained neglecting temporarily such gradients. Finally, when stirring is also interrupted (period III) the liquid surface will quickly heat up to an 'equilibrium' temperature as surface 2 continues to cool and surfaces 3 and 4 continue to heat till all four surfaces come to equilibrium. A straightforward lumped parameter model was devised[7] to describe this sequence of events quantitatively. The four unsteady-state energy balances include heat losses to environment, steam heating of surface 2 and a diffuse-grey formulation for radiative exchange among all four surfaces. During period III the liquid surface temperature is evaluated by assuming a conduction controlled the thermal boundary layer. Note that during distillation the vapour pressure is too low to produce significant inter-surface shielding. Similarly, during periods II and III the air presence in the gas space precludes such interference. The system was integrated numerically and the results of a representative calculation are shown in Fig. 1. Note that at the inception of period III only a small portion, if any (as actual cessation of all liquid motions, as assumed in the analysis, will occur in actuality well into period III), of the wall is around the threshold of 230 °C, while the liquid surface temperature remains well below this limit. In fact, the results are not very sensitive to parameter variations (such as the wall emissivity) always yielding surface liquid temperatures in the 200–220 °C range.

Figure 1 shows that the low liquid thermal diffusivity yields an isolation effect, whereby the liquid surface heats to nearly an equilibrium value consistent with transmitting a very low-energy flux. On the other hand, radiative exchange within surface 2 will tend to sharpen the temperature gradient near its boundary with surface 1 (I_1). The stainless wall has a thermal diffusivity two orders of magnitude higher than that of the liquid. One would expect significant conduction along the wall, creation of horizontal density gradients in the liquid and establishment of convective currents[8]. There is no direct information published about this complex heat transfer process but large-scale circulation currents are expected. These currents could destroy the thermal stratification picture developed above yielding even larger margins of maximum liquid temperatures from the 230 °C limit. We will argue instead that this convection–conduction phenomen acts in concert with the radiation process to produce an extremely localized hot stratum at the liquid surface.

As a simple first-order analysis we may view the vessel wall as a submerged fin. If T_w is the temperature of the fin base (at the liquid surface) and k_w the wall thermal conductivity, a quasi-static analysis indicates that the energy dissipated, by natural convection in a fluid of temperature T_∞, is given by $q = (Nu_B k_b/k_w)^{1/2} k_w(T_w - T_\infty)/B$. The Nusselt number is based on the fluid thermal conductivity k_b and fin thickness B. For the present situation it is easy to see that the square-root term is of the order of one. The fin-internal temperature gradient is, therefore, confined within a region of order B from the fin base. The same conclusion is also obtained for the case of a stratified liquid layer, that is with an imposed vertical temperature gradient in the bulk. An experimental confirmation of this trend was also obtained. A portion of metal (carbon steel) slab was immersed into a liquid (water) filled container. A temperature gradient in the well insulated non-submerged portion was established by steadily heating at one end and conducting into the liquid at the other. Determination of the resultant temperature fields was made by thermocouple measurements at a large number of points. One set of typical results is presented in Fig. 2. In particular note the coincidence of the fin and surrounding medium thermal boundary layers. Evaluation of growth histories of these thermal boundary layers indicates a process controlled by the liquid thermal diffusivity, that is $\delta \sim (\alpha_l t)^{1/2}$.

The combined effects of radiation and fin processes were also studied in an experiment more closely simulating the Seveso reactor using a cylindrical stainless-steel vessel, 76.2 cm in diameter and 122 cm tall. The wall thickness was selected at 0.277 cm such that the liquid surface area to wall cross-section area will approximate the Seveso reactor value of 60. This vessel was uniformly wrapped with cable heater and insulated to a degree approximating that of the Seveso reactor. The vessel was uniformly heated, while empty, to ~80 °C above room temperature (see Fig. 1 at beginning of period III). Room-temperature ethylene glycol was quietly admitted through a line

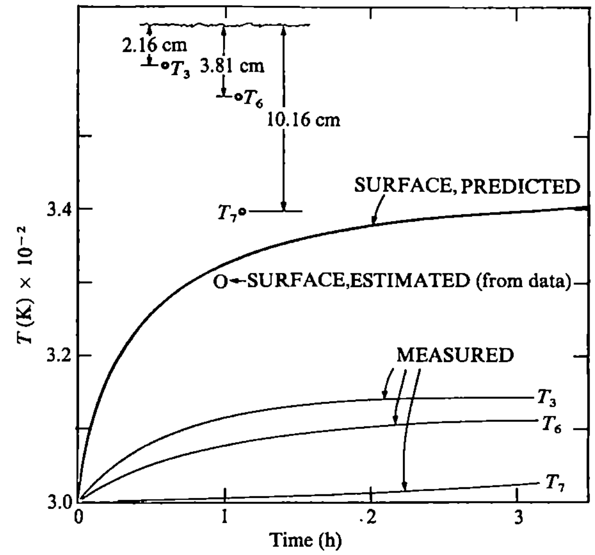


Fig. 3 Measured temperature transients near the liquid surface as indicated. Experimental and analytical simulation of the thermal transient in the Seveso reactor. The prediction is based on the lumped parameter radiation/diffusion model (emissivity $\varepsilon = 0.25$).

at the vessel base to a height of ~30 cm. The system was left at rest and temperature transients at the vessel walls and liquid contents were recorded with thermocouples. Due to space limitations only the essence of the results as summarized in Fig. 3 may be discussed here. The experimental data clearly show the downward development of the thermal boundary layer and they are in quantitative agreement with the $(\alpha_l t)^{1/2}$ dependence. The spatial dependence of these profiles also agrees with that predicted for transient conduction into a semi-infinite domain with a constant temperature of the surface. Hence, this surface temperature may be estimated and was found to be in good agreement with our theoretical predictions as shown in Fig. 3. A rapid liquid surface temperature increase of 30–40 °C similar both in trend as well as in magnitude to that calculated for the Seveso reactor (Fig. 1) is thus deduced demonstrating the nature of the stratification phenomena and the simulation capability of our experimental setup.

I thank Dr M. Saito for experimental assistance

Received 9 March; accepted 14 April 1981.

1. Hay, A. *Disasters* (to be published).
2. Hay, A. *Nature* **262,** 636 (1976); **267,** 384 (1977); **273,** 583 (1978).
3. Milnes, M. H. *Nature* **232,** 395 (1971).
4. Salomon, Ch. M. *Bericht uber den Störfall in Seveso, Sonderkolloquium am* 4 (Dechema-Haus, Frankfurt am Main, 1980).
5. Salomon, Ch. M. *Meet. Industrial Health and Safety Group of Soc. Chem. Ind.* London (1981).
6. Merzalanov, A. G. & Arerson, A. E. *Combust. Flame* **16,** 89 (1971).
7. Theofanous, T. G. *Meet. Industrial Health and Safety Group of Soc. Chem. Ind.* London (1981).
8. Theofanous, T. G. *Bericht uber den Störfall in Seveso, Sonderkolloquium am* 4 (Dechema-Haus, Frankfurt am Main, 1980).

Proterozoic magnetic overprinting of Archaean rocks in the Canadian Superior Province

Larry D. Schutts & David J. Dunlop

Department of Geology and Geophysics Division, Department of Physics, University of Toronto, Toronto, Canada M5S 1A7

We have measured the various components of natural remanent magnetization (NRM) exhibited by mafic extrusive and intrusive rocks from the Abitibi Subprovince of the Canadian Precambrian Shield. In addition to providing what seems to be a primary Archaean pole for the region, the results testify to three episodes of magnetic 'overprinting' at approximately 2,600, 2,100 and 1,800 Myr, respectively. The two earlier episodes coincide with the intrusion of Matachewan and Abitibi diabase dyke swarms, here and elsewhere in the Superior Province. Magnetic overprinting, in these cases, extends far beyond areas of surficial dyke exposure. The third episode postdates magmatic activity in the Abitibi region, but coincides in age with the 1,900–1,700 Myr Hudsonian orogeny in the Churchill and Bear Provinces and the Penokean orogeny in the Southern Province. Although magnetic overprints presumed to be of this age were known previously from Abitibi dykes[1] and Nipissing sills[2–4], our data are the first evidence of a 1,800 Myr magnetic overprinting on a regional scale in the Superior Province.

Palaeomagnetism is a powerful method of detecting and dating regional geological events in metamorphic terrains, particularly in Precambrian regions where complex structures, deep erosion, and various generations of burial and intrusion often confuse relations. Each geological event can result in thermal and/or chemical resetting of part of the pre-existing NRM. In favourable circumstances, careful multicomponent palaeomagnetic analysis can isolate vectorially superimposed NRM components corresponding to primary magnetization and one or more generations of NRM overprinting[5–8]. The geological events responsible are dated by comparing the palaeomagnetic pole corresponding to each NRM component with an established apparent polar wander path (APWP) for the continent or continental block.

The Abitibi greenstone belt of the Superior Province is particularly suited to a regional multicomponent palaeomagnetic study. The late Archaean and early Proterozoic units have been thoroughly mapped and the Laurentian APWP for this time interval is reasonably well established (Tracks 4, 5 and 6 of refs 9, 10). The Abitibi belt is characterized by mafic to felsic volcanic piles with coeval intrusions, volcanogenic sediments and intruding granitic batholiths[11,12]. In the Kirkland Lake, Ontario area, two such Archaean piles, plus the younger portion of a third pile, are preserved in a large east-plunging synclinorium[13]. The gabbroic intrusive bodies occur as both conformable sills[14] and as stocks, which seem to have been feeders for some of the flows[15]. Granitic bodies, apparently representing a later Archaean event, intrude the mafics on the north, west and south as batholiths and within the synclinorium as occasional stocks and plugs. Magmatic activity continued with intrusion of the north-south trending Matachewan diabase dyke swarm and essentially concluded with intrusion of the north-east trending Abitibi diabase dykes.

Archaean activity in the area was evidently episodic in occurrence. The youngest volcanic pile, which represents more than 70% of the 50,000 m total pile thickness[13], is bracketed in age by 2,710±2 Myr and 2,703±2 Myr U-Pb zircon dates[16]. These figures compare with 2,725±2 Myr and 2,703±3 Myr zircon ages bracketing the correlative sequence near Timmins, some 120 km to the west[17,18]. Granite ages are less narrowly defined. The oldest dates, determined by both K/Ar and Rb/Sr methods, fall within the 2,500–2,600 Myr range for the Round Lake Batholith[19,20]. The granites are cut by Matachewan dykes which, with a $\lambda(^{87}\mathrm{Rb})$ of 1.42×10^{-11} yr^{-1} (ref. 21), yield Rb/Sr isochron data implying an age of 2,633±93 Myr (ref. 22). The relative age relation, however, is clear. The Matachewan dykes are seen to intrude the granite in many locations and absolute ages for both units probably fall within the 2,600–2,700 Myr span. The time of Abitibi dyke intrusion is more precisely defined and implies a discrete early Proterozoic event at 2,150 Myr (refs 22, 23).

For our study, 157 block samples were taken over an areal extent of 5,200 km^2 in the Kirkland Lake Matheson area. The suite of samples includes basalts and gabbros from the youngest synclinorium sequence, as well as diabases from 15 Matachewan and three Abitibi dykes. Several 2.5-cm cylindrical specimens were cut from each sample and subjected to detailed stepwise thermal (to 700 °C) and alternating field (to 100mT or 1,000 Oe) demagnetization, using standard commercial demagnetizers. NRM intensities and directions were measured on a computerized commercial spinner magnetometer, which provided a minumum detectable signal of 10^{-4} A m^{-1} (10^{-7} e.m.u. cm^{-3}) and on-line statistical analysis of results. (A more detailed analysis of the alternating field and thermal demagnetization results, and the results of experiments in progress designed to characterize more fully the magnetic remanence carriers, will be published elsewhere[24].)

Except for a few basalt and gabbro samples, in which an intermediate (~350 °C) Curie temperature carrier of magnetization was indicated, all samples exhibited maximum thermal demagnetization blocking temperatures near 580 °C, the Curie temperature of pure (titanium-free) magnetite. Basalt, gabbro and Abitibi dyke blocking temperatures tended to be distributed from 100 or 200 °C to 580 °C, while Matachewan dyke blocking temperatures were generally concentrated just below 580 °C. This evidence of relatively pristine magnetite as the principal NRM carrier is in accord with petrographical observations from this and other studies[15,25] implying only very low-grade regional metamorphism in the area.

Different methods were used to analyse directional changes during demagnetization. Stable endpoints, Briden Index criteria[26], Zijderveld-type plots[27], vector subtraction diagrams[28]

Table 1 Summary of site mean palaeomagnetic results

	D	*I*	α_{95}	*k*	*R*	*N*	Palaeopole	
Characteristic directions								
Archaean basalt-gabbro	114°	−32°	11°	15	12.2	13	29° S	4° E
Matachewan dykes and baked contacts	21°	+13°	6°	44	13.7	14	45° N	70° E
Abitibi dykes and baked contacts	277°	+70°	8°	52	7.9	8	41° N	230° E
Secondary directions								
Archaean basalt-gabbro Matachewan-like	28°	+25°	11°	19	9.5	10	48° N	57° E
Archaean basalt-gabbro Abitibi-like	298°	+77°	9°	30	10.7	11	54° N	241° E
All units (sample averaged) post-Abitibi	166°	+56°	8°	14	26.1	28	4° S	291° E

D, *I* are declination and inclination of magnetization; α_{95}, *k* and *R* are Fisher statistical parameters; *N* is the number of sites averaged.

and converging remagnetization circles[29] all provided data relevant to the segregation of magnetic components. The isolated components were classified as either characteristic or secondary remanences. We define the 'characteristic NRM' in a multicomponent sum to be that component possessing the higher coercivities and blocking temperatures. A 'secondary NRM' is that component of lower coercivity and blocking temperature which overprints a characteristic NRM. Because our results imply that the reset magnetizations are of thermal as opposed to chemical origin, this scheme seems justified.

A standard three-tier (specimen/sample/site) analysis of characteristic and secondary NRM directions was carried out, a site being defined as a single stock, dyke, sill or flow and its baked contacts. Site-mean average magnetization directions are listed in Table 1 and plotted in Fig. 1. All NRM directions listed in Table 1 exhibit both normal and reverse polarities. Palaeopoles corresponding to each magnetization direction are shown in Fig. 2 and compared with the Tracks 5 and 6 of the North American APWP from refs 9, 10. Archaean ages have been updated to agree with the recent radiometric dates for the volcanics.

Abitibi and Matachewan dyke characteristic directions are well-defined and closely agree with the results of previous studies[1,10,30,31]. Evidence that these directions record the geomagnetic field direction at the time of dyke intrusion is provided by detailed Abitibi–Matachewan and Matachewan–gabbro contact tests in Munro Township. The pattern is the same in either case. Near the contact, NRM directions in the older unit have been reset parallel to those of the intrusive body. Well away from the contact, NRM directions in the older unit approach the characteristic direction. In an intermediate 'hybrid' zone, both NRMs are superimposed, the resultant direction swinging towards the older characteristic during magnetic cleaning. The maximum blocking temperature range of the overprinted NRM decreases in a regular fashion with distance from the contact. These data are strong evidence for partial to complete thermal resetting of primary NRM in the country rock at the time of dyke intrusion.

Basalt and gabbro samples were taken from the upper volcanic–plutonic pile of the synclinorium. The suite contains blocks from the Kinojevis tholeiite Group and the underlying Stoughton–Roquemaure and Larder Lake komatiite Groups. Three separate magnetic components, common to rocks from all three units, were isolated. Two of these components have the Abitibi and Matachewan dyke characteristic directions (Fig. 1*b*) and we believe them to be regional thermal overprints. The third component is similar to one found in other Superior Province Archaean units[10,32] and possibly records the geomagnetic field at the time these units were emplaced.

Several lines of evidence imply that the Abitibi-like and Matachewan-like components are secondary thermal overprints. They primarily affect the lower coercivity and blocking temperature spectra of the samples exhibiting them. They do persist throughout these spectra in some instances, but samples displaying this behaviour generally have coercivities and blocking temperatures centred about low absolute values. The directional changes with increasing demagnetization level are invariably in the sense suggested by the relative age data. Matachewan-like directions swing towards the older Archaean characteristic direction. Abitibi-like directions move towards the same Archaean direction or towards a Matachewan direction. These swings occur with both alternating field and thermal demagnetization. As both secondary and characteristic remanent magnetizations are lost by ~580 °C, the possibility of chemical overprinting seems remote. Partial thermal resetting of NRM as a result of, first, Matachewan and then Abitibi intrusion

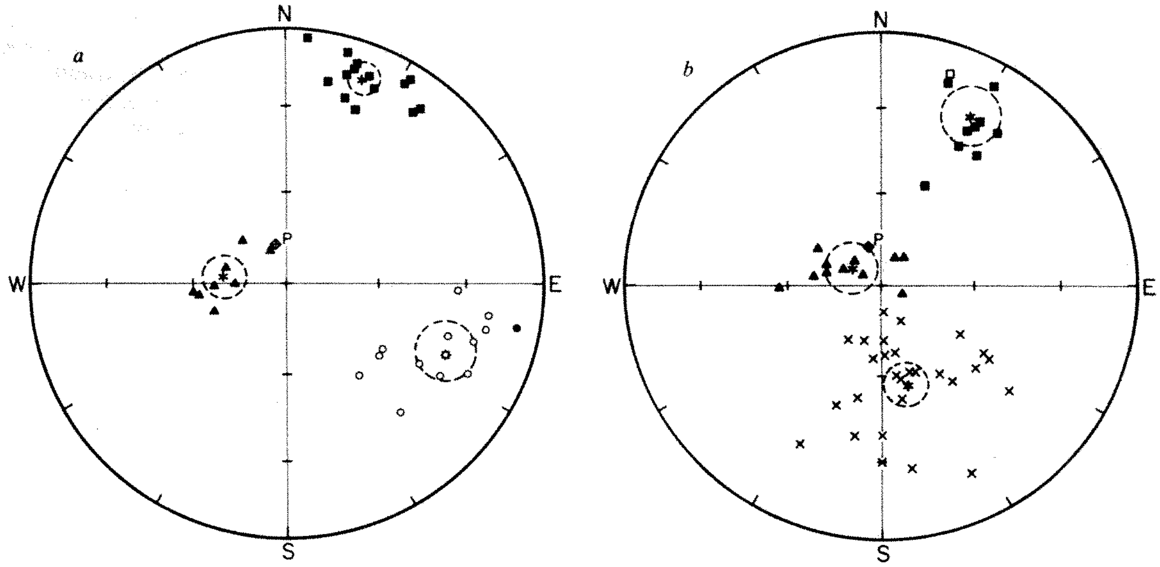


Fig. 1 Abitibi Subprovince site mean magnetization directions after optimum AF and/or thermal cleaning. Solid (open) symbols denote positive (negative) inclination. P (⊕) is the present Earth's field direction. ⋇, Average of site mean directions. *a*, Characteristic directions: ○, basalt–gabbro; ■, Matachewan diabase; ▲, Abitibi diabase. *b*, Secondary directions: ■, Matachewan-like; ▲, Abitibi-like; ×, Post-Abitibi (sample directions).

seems to be the most likely explanation for the occurrence of both secondary components. As the overprints are generally found in areas containing little or no surficial expression of such dyke intrusion, the entire region was probably reheated, although the existence of dykes which do not break the present surface is also possible.

Averages of the basalt and the gabbro site-mean characteristic directions are nearly identical. Hence, results for both rock types are combined to give a 'basalt-gabbro' characteristic direction. Throughout the area sampled, the gabbros occur primarily as bodies intruding the highly folded basalts. The basalt directions, corrected for structure, become essentially random and a case might be made for regional resetting at the time of gabbro intrusion. The common characteristic direction, however, is also found in basalts far removed from any obvious intrusion. It is also found in a conformable gabbroic sill within the volcanics, which yields a negative fold test (McCool Township). These observations, combined with geological suggestions that the volcanics and plutonics are consanguineous[15] and that the structure in the volcanics developed during deposition[33], provide strong evidence for the primary nature of the 'basalt-gabbro' characteristic.

This scheme does not take into account the ubiquitous granite intrusions. They apparently represent an Archaean event which postdates synclinorium mafic activity, but predates Matachewan dyke intrusion. Current palaeomagnetic evidence does not rule out the basalt-gabbro characteristic direction being the result of a regional granitic intrusion heating. Low granite intrusion temperatures and gravity data suggesting that the larger bodies are limited to the synclinorium periphery[11,34] oppose this possibility. Ultimate resolution of the question, however, awaits detailed palaeomagnetic study of the granitic units.

Although plutonic and volcanic activity in the area essentially ceased with intrusion of the Abitibi dykes, a post-Abitibi, thermally-reset NRM component has been isolated in the lower coercivity-blocking temperature ranges of 28 of the samples studied. These components partially overprint each of the other characteristic and secondary directions already discussed and show no preference for rock type or geographic locality. They are, however, encountered only sporadically throughout the varous sites. As a result, sample means have been used to calculate the average direction and palaeopole.

Poles corresponding to these sample-mean directions are grouped around the 1,700–1,900 Myr segment of Tracks 4 and 5 (Fig. 2). Such components have previously been found in both Abitibi dyke[1] and Nipissing sill[2,3] samples from the region and have been assigned secondary magnetization status and late Aphebian age on the basis of palaeomagnetic work in the Cobalt area[4]. By using the ^{40}Ar/^{39}Ar step-heating variant of K/Ar dating, Hanes and York have uncovered evidence for a late Aphebian heating event in Munro Township[23]. Samples taken from an Abitibi dyke yielded a 2,150 Myr age for the dyke interior and 1,700–1,800 Myr ages for the edge zone contact with a Matachewan dyke. Several samples from this Matachewan dyke and one from the country rock gabbro also gave 1,700–1,800 Myr plateaus. The detailed Abitibi–Matachewan contact test made for this study was conducted where the same Abitibi dyke investigated by Hanes and York cut a Matachewan dyke 150 m east of the one they sampled. The post-Abitibi secondary direction was found in both units, most frequently near the contact. A late Aphebian age for this post-Abitibi NRM component is strongly implied.

More than half of the post-Abitibi directions discovered occurred in samples taken near intrusive contacts or zones of some faulting. A high heat flow in the region 1,700–1,900 Myr ago, facilitated by hot fluid transport through faults and intrusive contact zones of weakness, could explain both magnetic and argon resetting. A re-echoing of the Southern Province Penokean orogeny, for example, might be postulated as a source for the high heat flow. Although late Aphebian chemical resetting of remanence in some samples cannot be ruled out, Curie temperatures and preliminary microscopic examinations in reflected light suggest that thermal resetting should be the favoured hypothesis. Single Curie temperatures in the range 570–580 °C and intermediate deuteric alteration states with little sign of maghaemitization or haematization of the titanomagnetite phase imply no significant metasomatic alteration of the remanence carrier since formation.

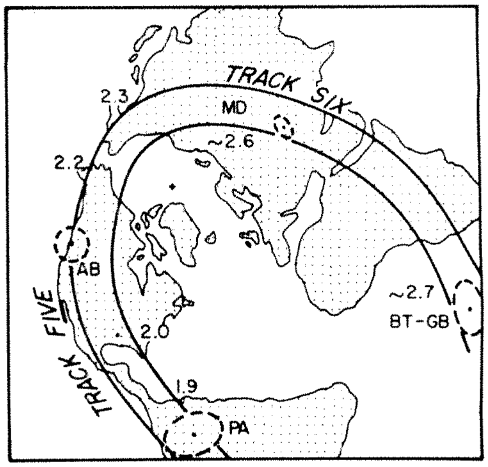


Fig. 2 Palaeomagnetic poles corresponding to Fig. 1 NRM directions as compared with the Laurentian APWP for the interval 1,900–2,700 Myr. BT-GB, basalt-gabbro characteristic. MD, Matachewan diabase characteristic—secondary. AB, Abitibi diabase characteristic—secondary. PA, post-Abitibi secondary.

Alternatively, the resetting may not be related to faulting and intrusive contacts on a regional scale and a late Aphebian unroofing event might be invoked to explain the younger ages and directions. The region was covered by Huronian sediments at ~2,300 Myr (ref. 35). A Schwarz-type[36,37] depth of burial calculation, using minimum blocking temperatures of the pre-overprint component, has been made for samples from the Munro Township Abitibi–Matachewan contact. The results imply an ambient temperature for the present erosional surface of 190 °C at the time of Abitibi dyke intrusion. Assuming a 25 °C km^{-1} geothermal gradient, this temperature implies a 7–8 km depth of burial at 2,150 Myr. The area was still covered, then, during the middle Aphebian, but a late Aphebian unroofing may certainly be postulated.

The close agreement of our independently dated post-Abitibi palaeopole for the Superior Province with palaeopoles ~1,800 Myr in age from the Churchill and Bear Provinces has broad implications for the tectonic history of Laurentia. This evidence implies that the Churchill orogen and the adjacent, relatively stable Superior craton occupied their present relative positions during the 1,900–1,700 Myr Hudsonian orogeny. The orogeny, therefore, may not have been the result of large-scale relative motions between these two blocks, such as are involved in modern day plate collision and subduction. The Hudsonian orogeny may represent an essentially ensialic reactivation of the Archaean Laurentian craton that incorporated both the Superior and Churchill Provinces.

Our preliminary Abitibi Subprovince results demonstrate that the geological history of complex Archaean terrains can be clarified with the help of palaeomagnetism. Resolved multi-component magnetizations can provide, in effect, a capsule history of thermal events throughout an entire region.

This research has been supported by Natural Sciences and Engineering Research Council Canada and Energy Mines and Resources Canada Research Grants to D. J. D. and by University of Toronto and Ontario Graduate Fellowships to L. D. S.

Received 22 December 1980; accepted 8 April 1981.

1. Larochelle, A. *Can. J. Earth Sci.* **3**, 671 (1966).
2. Patel, J. P. & Palmer, H. C. *Can. J. Earth Sci.* **11**, 353 (1974).
3. Symons, D. T. A. & Londry, J. W. *Can. J. Earth Sci.* **12**, 940 (1975).
4. Roy, J. L. Lapointe, P. L. *Can. J. Earth Sci.* **13**, 749 (1976).
5. Buchan, K. L., Berger, G. W., McWilliams, M. O., York, D. & Dunlop, D. J. *J. Geomagn. Geoelectr.* **29**, 401 (1977).
6. McElhinny, M. W. & Senanayake, W. E. *J. geophys. Res.* **85**, 3523 (1980).
7. Beckmann, G. E. J. & Mitchell, J. G. *Earth planet. Sci. Lett.* **30**, 269 (1976).

8. Morgan, G. E. *Nature* **259,** 382 (1976).
9. Irving, E. & McGlynn, J. C. *Trans. R. Soc.* **A280,** 433 (1976).
10. Irving, E. & Naldrett, A. J. *J. Geol.* **85,** 157 (1977).
11. Goodwin, A. M. & Ridler, R. H. *Geol. Surv. Canad. Pap.* 70–40, 1 (1970).
12. Goodwin, A. M. *Geol. Ass. Can. Spec. Pap.* **16,** 205 (1977).
13. Jensen, L. S. *Ont. Geol. Surv. Misc. Pap* No. 90, 64 (1979).
14. Pyke, D. R., Naldrett, A. J. & Eckstrand, O. R. *Bull. geol. Soc. Am.* **84,** 955 (1973).
15. Jensen, L. S. *Ont. Geol. Surv. Rep.* 165 (1978).
16. Nunes, P. D. & Jensen, L. S. *Ont. Geol. Surv. Misc. Pap.* 92 (1980).
17. Pyke, D. R. *Ont. Geol. Surv. Misc. Pap.* No. 82, 73 (1978).
18. Nunes, P. D. & Pyke, D. R. *Ont. Geol. Surv. Misc. Pap.* 92 (1980).
19. Aldrich, L. T. & Wetherill, G. W. *J. geophys. Res.* **65,** 337 (1960).
20. Purdy, J. W. & York, D. *Can. J. Earth Sci.* **5,** 699 (1968).
21. Steiger, R. H. & Jäger, E. *Earth planet. Sci. Lett.* **36,** 359 (1977).
22. Gates, T. M. & Hurley, P. M. *Can. J. Earth Sci.* **10,** 900 (1973).
23. Hanes, J. A. & York, D. *Can. J. Earth Sci.* **16,** 1060 (1979).
24. Schutts, L. D. (in preparation).
25. Jolly, W. T. *Can. Miner* **12,** 499 (1974).
26. Briden, J. C. *J. geophys. Res.* **77,** 1401 (1972).
27. Zijderveld, J. D. A. in *Methods of Paleomagnetism* (eds Collinson, D. S., Creer, K. M. & Runcorn, S. K.) 172 (Elsevier, Amsterdam, 1967).
28. Halls, H. C. *Earth planet. Sci. Lett.* **43,** 303 (1979).
29. Halls, H. C. *Phys. Earth planet. Inter.* **16,** 1 (1978).
30. Strangway, D. W. *J. Geol.* **72,** 648 (1964).
31. Fahrig, W. F., Gaucher, E. H. & Larochelle, A. *Can. J. Earth Sci.* **2,** 278 (1965).
32. Dunlop, D. J. *Phil. Trans. R. Soc.* (in the press).
33. Jensen, L. S. *Toronto '78 Field Trips Guidebook,* 237 (Geological Association of Canada, 1978).
34. Middleton, R. S. *Ont. Div. Mines, Geosci. Rep.* 135 (1976).
35. Fairbairn, H. W., Hurley, P. M., Card, K. D. & Knight, C. J. *Can. J. Earth Sci.* **6,** 489 (1969).
36. Schwarz, E. J. *Can. J. Earth Sci.* **14,** 82 (1977).
37. Buchan, K. L. & Schwarz, E. J. *Can. J. Earth Sci.* (submitted).

Low-angle normal faults in the Basin and Range Province: nappe tectonics in an extending orogen

Brian Wernicke

Department of Earth and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Cenozoic normal fault mosaics bounded beneath by a basal fault in the Basin and Range Province (BRP) have traditionally been described either in terms of large-scale surficial gravity sliding or by some form of *in situ* lower plate accommodation. I suggest here that these areas may be an extensional analogue to thin-skinned compressional orogens, a process which may even dominate active BRP tectonics.

These terranes consist of a mosaic of rock sheets bounded by high- and low-angle normal faults[1,2] which resemble (and have been misinterpreted as) compressional nappe piles. In many areas, the normal faults 'bottom out' into a single, subhorizontal fault zone below which either minor or no tectonism has occurred[3–6], or a roughly synchronous, more ductile style of deformation is present[7,8]. The minimum areal extent of the basal faults is typically measured in thousands of square kilometres. For example, in east-central Nevada, normal fault systems developed in Palaeozoic miogeoclinal strata, Mesozoic and Tertiary intrusives, and Tertiary volcanics end abruptly at a basal fault (Snake Range decollement[3], Fig. 1) accompanied by a thin (0–100 m) zone of tectonite marble. Offsets on the basal and higher faults are difficult to contrain, but Nelson reports a 20-km minimum horizontal offset for one fault alone in the Deep Creek Range[9]. My palinspastic reconstruction of the Northern Egan and Cherry Creek ranges (Fig. 2) indicates a 16.5-km offset for one sheet in the mosaic, and the per cent increase in original width of the area is ~300%. As the basal and higher faults in Fig. 1 are commonly subparallel to bedding and often cut out thousands of metres of section, total offsets are probably at least in the range of several tens of kilometres. Most authors except one[10] who have mapped in the area outlined in Fig. 1 agree that the transport direction of the allochthonous sheets was eastwards. The faults also tend to cut consistently downsection to the east[11,12], (Fig. 2). This terrane apparently has every attribute of a thin-skinned compressional orogen (basal decollement, far-travelled allochthons, consistent direction of section transgression and transport) except (1) the low-angle faults habitually omit rather than repeat section, (2) associated high-angle faults are predominantly normal instead of reverse, and (3) folding is of subdued importance[13]. Though early workers attempted to correlate these faults with Mesozoic thrusting in the Sevier belt to the east[3], there is considerable evidence to support a Tertiary, extension-related origin for most of the deformation[13–15].

Models for these areas must account for three observations: (1) the basal and higher faults predominantly represent extension; (2) the rocks below the basal faults are tectonically inert during transport of the allochthonous sheets and high-angle fault blocks; and (3) the areal extent and transport distances of the allochthons are large. Current models fall into two categories. The first is a localized gravity-slide model in which domal upwarps shed their cover in a megalandslide fashion[15–18] (Fig. 3*a*). The second invokes penetrative ductile stretching and/or igneous intrusion to drive extension in the immediately overlying brittle fault mosaic[2,7,8,19,20] (Fig. 3*b*). Landslide models are weakened by the scarcity of geometrically required areas of shortening correlative in age, size and direction of transport with the extensional terranes. Thus, the model agrees with observations (2) and (3), but not with observation (1). *In situ* ductile stretching or intrusion models lose appeal when it is noted that most metamorphic fabrics and igneous bodies (if any) beneath the basal faults are of inappropriate age or geometry to be rigorously linked to normal faulting[5,6,21,22]. Thus, the model is in harmony with observations (1) and (3), but not with observation (2).

The key to the problem may lie in our understanding of compressional orogens, in which the horizontal movement of thin sheets of rock above an underthrusting slab may occur many

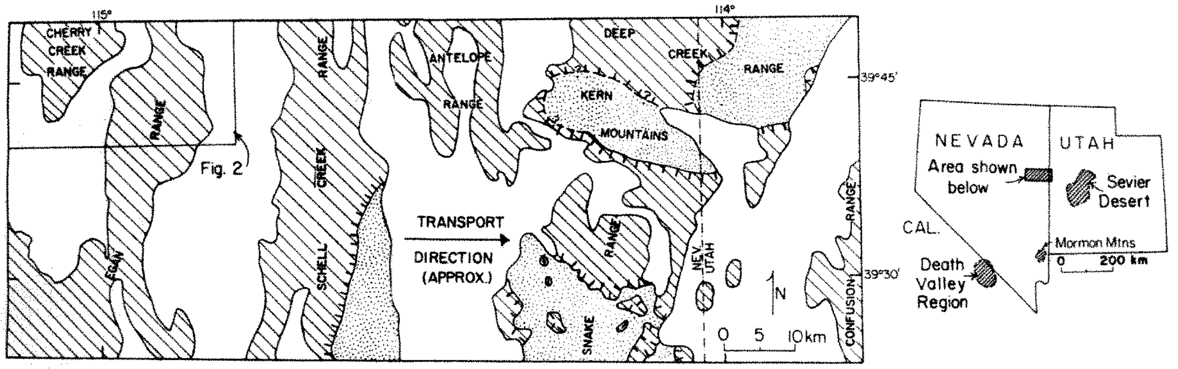


Fig. 1 Location map and tectonic interpretation of east-central Nevada, showing autochthon (stippled), basal fault zone (line with tick marks, queried where uncertain) and normal fault mosaic (hatched). Basal fault in the Schell Creek Range is locally tilted steeply westwards, implying continued deformation after it ceased moving, possibly on a lower basal fault. Tilting and faulting in the Northern Egan Range may be related to this second event rather than to movement on the basal fault shown here.

0028–0836/81/260645—04$01.00

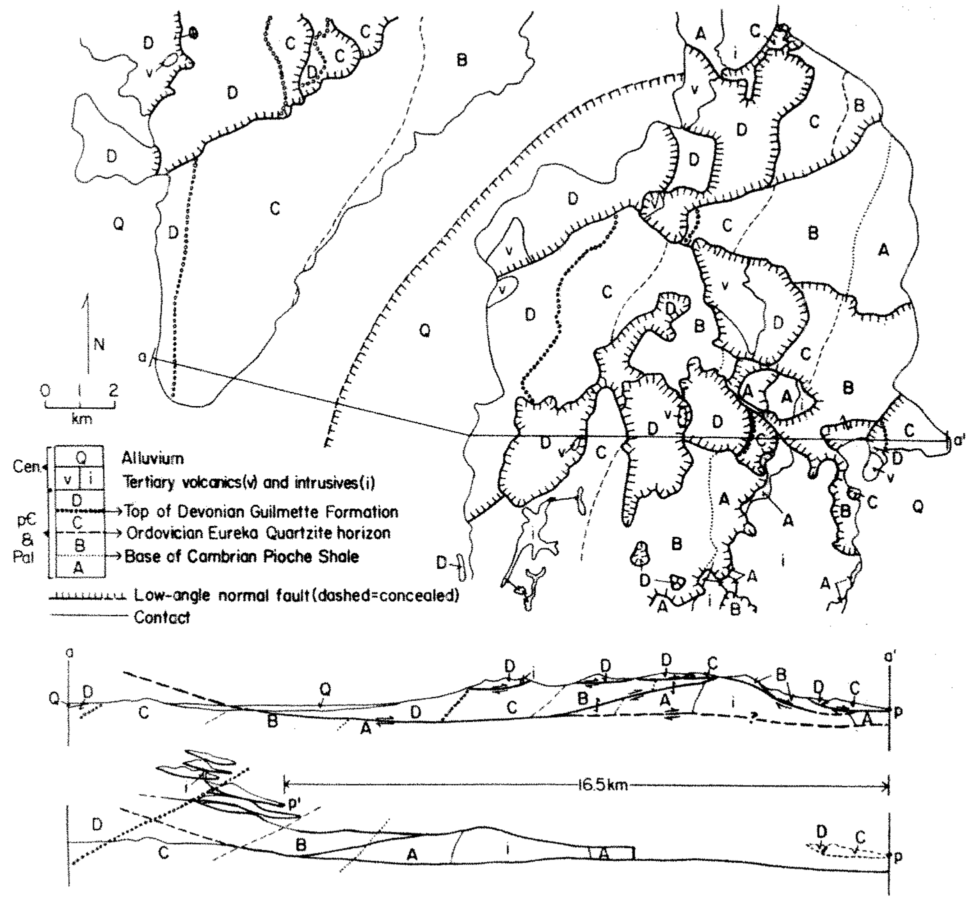

N
0 1 2
km
Cen.
Q Alluvium
v i Tertiary volcanics(v) and intrusives(i)
pC & Pal
D →Top of Devonian Guilmette Formation
C → Ordovician Eureka Quartzite horizon
B → Base of Cambrian Pioche Shale
A
Low-angle normal fault (dashed=concealed)
Contact
a
a'
p
p'
16.5km

Fig. 2 Highly simplified map, cross-section and palinspastic reconstruction of the area indexed in Fig. 1. Reconstruction does not remove tilting, which was probably synchronous with faulting. Armstrong[15] interpreted the sheets as having formed exclusively by high-angle normal faulting, then subsequently rotated to the horizontal. Geology by Fritz[11].

tens (even hundreds) of kilometres from the nearest possible site of coeval lower plate shortening. Applying this concept to extensional tectonics to explain relationships observed in the BRP, I suggest that the basal faults 'root' into the crust (and perhaps the mantle) at a low angle, thereby serving as a narrow zone of decoupling between two thin sheets of rock (Fig. 3*c*). Towards the thick end of the upper plate, extensional faulting may be negligible or absent, but towards the thin end it loses its ability to remain coherent and becomes a thin-skinned extensional fault terrane. Lower plate extension may therefore be far removed horizontally (perhaps a distance many times the areal width of upper plate normal faulting) and at much greater depth than the thin-skinned extensional belt. This hypothesis allows for a rigid autochthon and eliminates the need for immense terranes of coeval shortening, thus satisfying all three observations.

As shown in Fig. 3*c*, rock sheets (nappes[23]) may be carried along the basal fault and accreted to the autochthon, kinematically similar to nappe emplacement along the soles of overthrust masses in compressional belts. Note that the nappes in Fig. 3*c* have converged on one another, a relationship that superficially resembles shortening. However, the amount of movement between them is actually a measure of crustal divergence. If the faults had developed in a pristine sedimentary

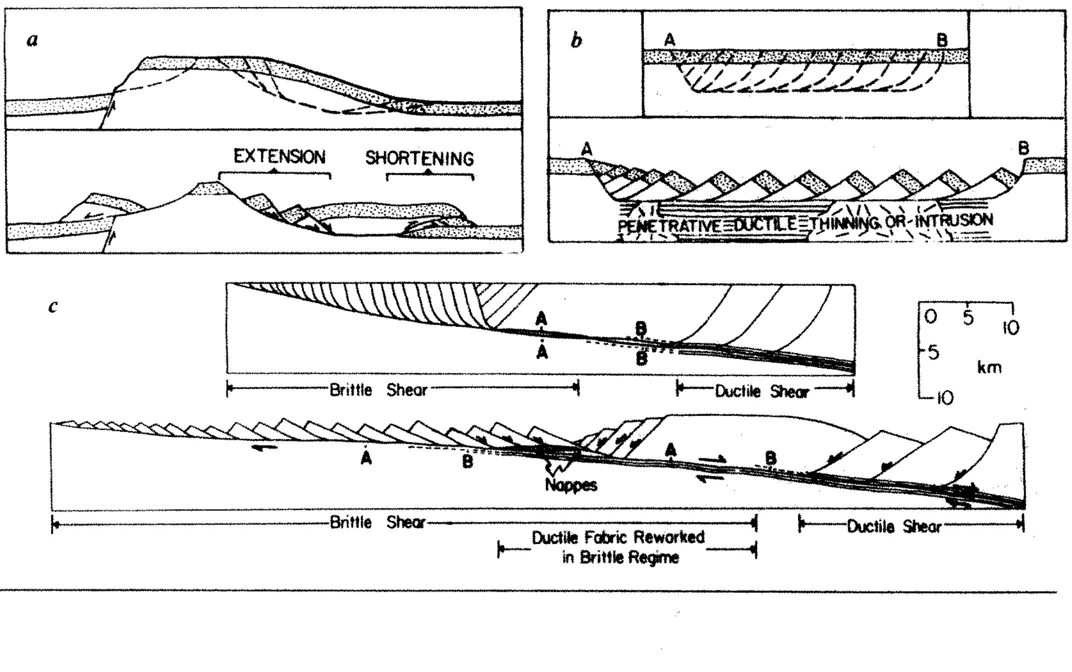

a
EXTENSION
SHORTENING
b
A
B
PENETRATIVE DUCTILE THINNING OR INTRUSION
c
A
B
Brittle Shear
Ductile Shear
0 5 10
5
10
km
Nappes
Ductile Fabric Reworked in Brittle Regime

Fig. 3 First-order kinematic schemes to generate low-angle normal fault mosaics above abasal fault. *a*, Megalandslide model; *b*, *in situ* ductile stretching or intrusion model; *c*, rooted, low-angle normal fault model.

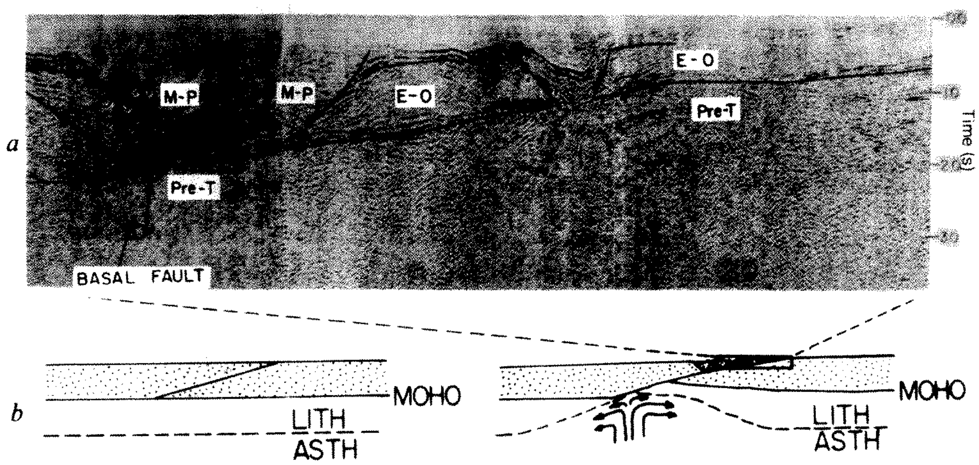


Fig. 4 *a*, North-east trending, 40-km long two-way time section through the Sevier Desert area[29]. Pre-T, pre-Tertiary; E–O, Eocene–Oligocene; M–P, Miocene–Pliocene. *b*, Lower crust–mantle extrapolation of faults such as that in *a*. Note that the geometry of the Moho is indistinguishable from how it would look if the crust had penetratively stretched. (Seismic line reproduced with permission of the Rocky Mountain Association of Geologists.)

sequence, the fault between the nappes would place younger rocks on top of older rocks, a direct consequence of the fault rooting in the same direction as transport. Had the same juxtaposition occurred along a fault which rooted opposite its transport direction, movement between the nappes would be an expression of true crustal convergence. Large, rooted low-angle normal faults may form zones of ductile shear along deeper segments. If the displacement is large enough, these zones may shallow and be tectonically reworked in brittle conditions[21]. This scheme of nappe emplacement, combined with superjacent 'domino-style'[24] and listric normal faulting, are major mechanisms of extension above the basal fault. Although some aspects of this model are not new[25,26], its application to low-angle normal fault terranes in the BRP has been neglected in favour of gravity and *in situ* models.

Distinction between low-angle normal fault terranes and active BRP tectonics has been strongly emphasized by some workers[27], and others have related the apparent difference in style (low-angle compared with high-angle) to plate tectonics[28]. However, the neotectonics of the BRP are apparently sufficiently ill-characterized at depth to entertain the hypothesis that rooted, low-angle normal faults are an important mode of Plio-Quaternary extension, perhaps the fundamental control of its famous topography. For example, the Amargosa Chaos[4], a striking example of low-angle normal faulting[1], involves mid- to late-Pliocene volcanics, and hence was active well within what some consider to be a province-wide period of high-angle faulting. Further evidence suggesting that the present mechanism may control active BRP tectonics comes from seismic reflection profiles of the Sevier Desert area[29]. Figure 4*a* and three other profiles reveal a minimum area of $\sim 5{,}000$ km^2 underlain by a gently west-dipping low-angle normal fault, above which Miocene–Pliocene basin fill and Quaternary basalt flows are offset by normal faults[29–31].

The basal fault in Fig. 4*a* has been interpreted as a reactivated thrust fault[29], a common interpretation of low-angle normal faults situated in older thrust terranes[25,32,33]. However, recent studies of some exhumed thin-skinned extensional areas suggest that pre-existing anisotropies are not a prerequisite for low-angle normal faults[5], and thrust faults, if present, need not control their geometry. In the Mormon Mountains of southern Nevada (Fig. 1), a Miocene low-angle normal fault event is superposed on a system of Mesozoic thrusts, a situation geometrically identical to that in the Sevier Desert[34]. The normal faults there truncate the thrusts at high angle, and clearly involve Precambrian crystalline rocks of the thrust autochthon. Based on this study, the normal faults in the Sevier Desert may be unrelated to Mesozoic thrusts prominent in the surface geology of the area.

Figures 3*c* and 4*a* suggest that conservative estimates of province-wide extension (10–30%) based on its proportionality to regional tilts of Tertiary strata may have no meaning[35], because large, horizontal translations of rock masses may occur without appreciable rotation of strata (as well as opposite tilts within the same allochthon). The present model therefore compliments larger estimates (50–100%) deduced by strike-slip fault reconstruction (work in preparation) and crustal thickness arguments[20].

Another critical aspect of the model is that it presents an alternative to penetrative ductile thinning or intrusion to accommodate extension at crustal (and mantle) levels not exposed in the BRP. Whereas igneous intrusion may increase the width of the crust, it cannot thin it; at best, it can retain its original thickness, if the intrusions are vertical dykes. Alternatively, it can actually thicken, if the intrusions are sill-like, a property of the large early- to mid-Tertiary batholiths of southern Arizona[22]. Because intrusion cannot therefore thin deeper crustal levels, it is difficult to envisage it as a major mechanism of BRP extension. Although taffy-like stretching of crustal dimensions does not encounter these geometrical difficulties, there is no compelling evidence to suggest that it is a major process at depth. An attractive alternative, consistent with surface processes, is that large, rooted normal faults extend at a low angle deep into the lithosphere, and that extension at the surface is due to discrete shear between large coherent sheets (Fig. 4*b*). Pull-apart may then ultimately be accommodated by asthenospheric convection rather than any form of stretching.

I thank colleagues at MIT for discussions. This work was supported by NSF grant EAR 7713637 awarded to B. C. Burchfiel.

Received 18 December 1980; accepted 17 March 1981.

1. Wright, L. A. & Troxel, B. W. In *Gravity and Tectonics* (eds de Jong, K. A. & Scholton) 397–407 (Wiley, New York, 1973).
2. Anderson, R. E. *Bull. geol. Soc. Am.* **82,** 43–58 (1971).
3. Misch, P. *Intermountain Ass. petrol. Geol. 11th a. Field Conf. Guidebook,* 17–42 (1960).
4. Noble, L. F. *Bull. geol. Soc. Am.* **52,** 941–1000 (1941).
5. Davis, G. A., Anderson, J. L., Frost, E. G. & Shackelford, T. J. *Geol. Soc. Am. Mem.* **153,** 79–129 (1980).
6. Compton, R. R., Todd, V. R., Zartman, R. E. & Naeser, C. W. *Bull. geol. Soc. Am.* **88,** 1237–1250 (1977).
7. Rehrig, W. A. & Reynolds, S. J. *Geol. Soc. Am. Mem.* **153,** 131–158 (1980).
8. Davis, G. H. *Geol. Soc. Am. Mem.* **153,** 35–78 (1980).
9. Nelson, R. B. *Bull. Am. Ass. petrol. Geol.* **53,** 307–339 (1969).
10. Fritz, W. H. *Nevada Bureau of Mines Map No.* 35 (1968).
11. Bick, K. F. *Bull. Utah Geol. miner. Soc.* **77** (1966).
12. Dechert, C. P. thesis, Univ. Washington (1967).
13. Young, J. C. *Intermountain Ass. petrol. Geol. 11th a. Field Conf. Guidebook,* 158–172 (1960).
14. Lee, D. E., Marvin, R. F., Stern, T. W. & Peterman, Z. E. *U. S. geol. Surv. Prof. Pap.* **700-D,** D92–D102 (1970).
15. Armstrong, R. L. *Bull. geol. Soc. Am.* **83,** 1729–1754 (1972).
16. Drewes, H., *U. S. geol. Surv. Prof. Pap.* **557** (1967).
17. Moores, E. M., Scott, R. B. & Lumsden, W. W. *Bull. geol. Soc. Am.* **79,** 1703–1726 (1968).
18. Seager, W. R. *Bull. geol. Soc. Am.* **81,** 1517–1538 (1971).
19. Proffett, J. M. Jr. *Bull. geol. Soc. Am.* **88,** 247–266 (1977).
20. Hamilton, W. B. *Pacif. Coast Paleogeogr. Symp.* **2,** 33–70 (1978).
21. Snoke, A. W. *Geol. Soc. Am. Mem.* **153,** 287–334 (1980).
22. Kieth, S. B. *et al. Geol. Soc. Am. Mem.* **153,** 217–268 (1980).
23. Gary, M., McAfee, R. Jr. & Wolf, C. L. (eds) *Glossary of Geology* (Americal Geological Institute, Falls Church, 1972).
24. Ransome, F. L., Emmons, W. H. & Garrey, G. H. *Bull. U. S. geol. Surv.* **407** (1910).
25. Bally, A. W., Gordy, P. L. & Stewart, G. A. *Bull. Can. petrol. Geol.* **14,** 337–381 (1966).
26. Hunt, C. B. & Mabey, D. R. *U. S. Geol. Surv. Prof. Pap.* **494-A** (1966).
27. Reynolds, S. J. & Rehrig, W. A. *Geol. Soc. Am. Mem.* **153,** 159–176 (1980).

28. Thompson, G. A. & Zoback, M. L. *Tectonophysics* **61,** 149–181 (1979).
29. McDonald, R. E. *Symp. on Cordilleran Hingeline,* 281–317 (Rocky Mountain Assocation of Petroleum Geologists, 1976).
30. Hoover, J. D. *Brigham Young Univ. Geol. Stud.* **21,** 3–72 (1974).
31. Clark, E. *Brigham Young Univ. Geol. Stud.* **24,** 87–114 (1976).
32. Dahlstrom, C. D. A. *Bull. Can. petrol. Geol.* **18,** 332–406 (1970).
33. Royse, F. Jr, Warner, M. A. & Reese, D. C. *Deep Drilling Frontiers in Central Rocky Mountains,* 41–54 (Rocky Mountain Association of Petroleum Geologists, 1975).
34. Wernicke, B. P. *Geol. Soc. Am. Abstr. Prog.* **13,** 113–114 (1981).
35. Stewart, J. H. *Bull. geol. Soc. Am.* **91,** 460–464 (1980).

Deep seismic soundings in India and the origin of continental crust

K. Roy Chowdhury & R. B. Hargraves*

National Geophysical Research Institute, Hyderabad-500 007, India

The present day continental crust beneath stable cratons is generally thought to be 30–50 km thick[1,2] although there are divergent views as to its initial thickness and the associated thermal and tectonic regimes[3,4]. Likewise, the lithosphere is believed to have been thinner in the Archaean, but its relationship to the crust at that time has not been specified. Structural details regarding the Archaean crust revealed by a Deep Seismic Sounding (DSS) study across the South Indian Craton[5], supplemented by similar data from the Ukrainian Shield[1], show a relationship between crust and mantle which apparently places important constraints on the thermo-tectonic conditions prevailing at the time of formation of the crust. We interpret these details as indicating that at one time the crust constituted the lithosphere, the Moho marking its base.

Figure 1 shows the main geological units of the Indian Peninsular Shield. Apart from narrow coastal strips of younger rocks, the terrain is entirely Precambrian in age. Radiometric data[6–9] indicate that the development of the South Indian Craton took place 3,400–2,000 Myr ago. Unmetamorphosed dyke-swarms are quite common in the Shield and one such swarm has been dated at 2,420±250 Myr (Rb/Sr whole rock isochron)[8]. The Champion Gneiss of Kolar and parts of the Closepet Granites give, however, whole rock and mineral ages around 2,000 Myr (ref. 7), indicating a major, later, thermal event. On the other hand, the rocks of the Cuddapah Basin are reported to be younger than 1,700 Myr (ref. 10). Thus it seems that stable craton was essentially established 2,000 Myr ago.

The resolving power of deep seismic sounding (DSS) developed in the Soviet Union and the more recent deep reflection profiling (DRP) used in the United States is at least an order of magnitude greater than of those based on earthquake or long-range refraction studies, because of the higher frequencies and the correspondingly smaller wavelengths being used to 'sound' the Earth. Whereas seismological investigations use wavelengths ranging from a few hundreds to a few tens of kilometres (free oscillations, surface wave—and body wave—studies), the DSS and the DRP techniques involve shorter waves only a few hundreds of metres long and less. Thus seismic discontinuities can be resolved within the crust to a degree much greater than before. For example, in the DRP technique a vertical resolution of 150 m and a horizontal resolution of 500 m has been reported[11].

In DSS[1,5] mostly wide angle reflections are recorded from discontinuities situated at depths down to and below Moho, the nominal crust–mantle boundary. Suitably selected shot points (typically 40 km apart) and continuous recording up to at least 200-km profile segments (using grouped geophones typically 200 m apart) enable overlapping and reciprocal coverage. The recorded waves include both reflections and refractions from shallow, intermediate and deep boundaries, and diffractions. The wide angle reflections from the Moho show up as high energy arrivals (due to the sudden and large increase in the reflection coefficient near the wide angle range resulting in almost total reflection) and dominate the records in the distance range of 80–150 km from the shot points.

The DSS data reveal that the crust is broken into blocks by deep-faults. In addition, the structure within the blocks themselves is complex, characterized by a large number of reflector and refractor segments that are up to a few tens of kilometres in length and exhibit irregular dips, even within individual fault-blocks, up to 30°–40°. The Moho segments are more continous with more regular dips. The simple three-layered crustal model (sediments, granite, gabbro/granulite) inherited from traditional seismology has consequently to be replaced by one that is more complex. The model must contain both vertical and lateral heterogeneities, possibly accompanied by many localized velocity reversals (ref. 1, p. 44, ref. 12).

Figure 2 shows the disposition of the Moho along the 600-km long Kavali–Udipi profile[5]. Here the crust is seen to be cut up into 17 major and minor blocks by deep and intermediate depth faults and thrusts. Many of these faults offset the Moho, up to as much as 7 km (fault 19, between blocks XIII and XIV). The depth to the Moho, in this section, ranges from 34 to 41 km and averages about 38 km. However, within each of these blocks the Moho is essentially horizontal, although short inclined segments are occasionally present.

As seismic profiling yields only apparent dips, inversion of DSS data in terms of the complete crustal structure would require a gridlike ground coverage, not yet available for the region considered here. However, the Kavali–Udipi profile, shown in Fig. 1 was deliberately oriented to be broadly perpendicular to the regional strike. Therefore the dips reported here are believed to be close to their true values. (In Fig. 2 the vertical scale is exaggerated 3:1, thereby accentuating apparent dips).

Figure 3 (no vertical exaggeration) shows the middle part of the Kavali–Udipi crustal section in some detail. The intra-

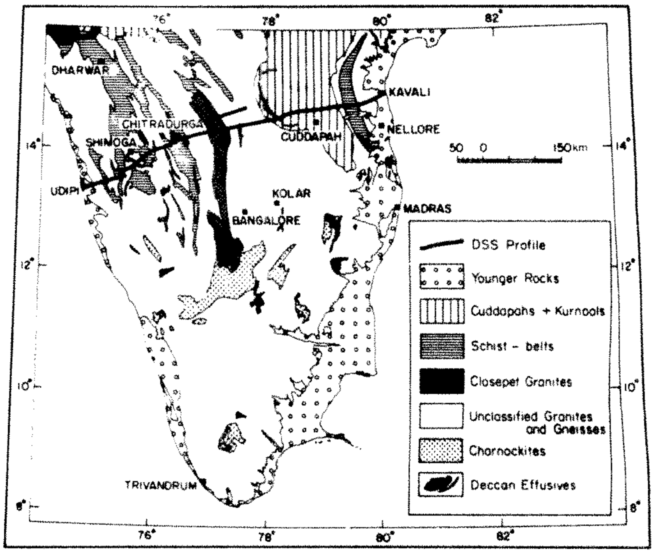


Fig. 1 The Indian Peninsular Region showing the location of the Kavali–Udipi DSS profile (recorded during 1972–75). The Shimoga and the Chitradurga schist-belts in the western part are predominantly sedimentary, rather than of mafic/volcanic type. Radiometric data[6,7] place them as late Archaean-early Proterozoic and due to their prominent marble–orthoquartzite association these have been termed 'greenstone-like geosynclinal piles'[18]. The greenstones proper, dominated by volcanics, are distributed widely across the peninsula, represented by the Nellore schist-belt on the east[18], and numerous remnants (such as Kolar) including a narrow patch on the eastern margin of the Chitradurga belt. The crescent-shaped Cuddapah Basin in the east contains mainly unmetamorphosed and minimally deformed platformal deposits of mid-Proterozoic age[10]. On its eastern flank the basin is truncated by a deep thrust[5] along which the Nellore schist-belt has overridden and overturned the Upper Cuddapahs. Near the western border of the Cuddapah Basin, abundant basaltic dykes and traps are exposed.

---

*Permanent address: Department of Geological and Geophysical Science, Princeton University, Princeton, New Jersey 08544, USA.

0028-0836/81/260648—03$01.00

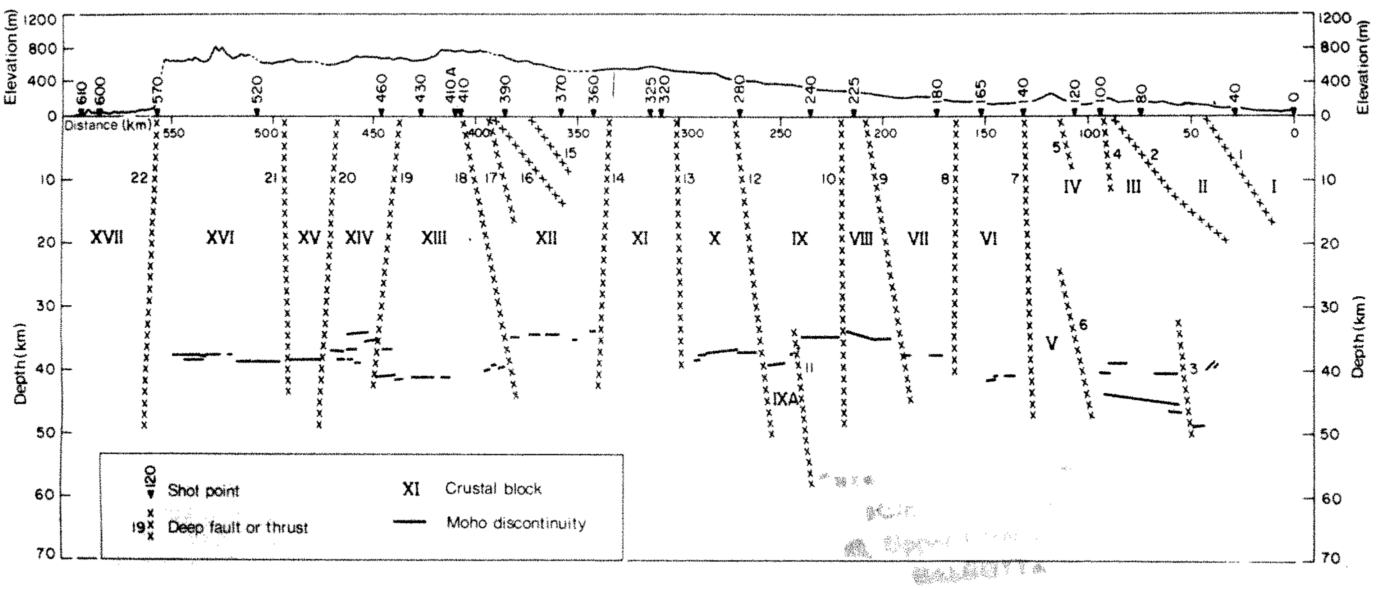


Fig. 2 Crustal thickness and block structure along the Kavali–Udipi DSS profile. As the vertical scale is exaggerated by 3:1 the dips are accentuated.

crustal reflector segments in blocks VI, VIII and X are fairly representative of the rest of the section shown in Fig. 2 and are characterized by rather irregular dips. These may be due to localized folding of crustal layers. This folding, however, has evidently not affected the underlying Moho, which is almost horizontal.

In blocks VII and IX, the fine structure is relatively consistent throughout, with dips up to 20°; yet, here too, the Moho is still essentially horizontal. In these two blocks in particular, the Moho seems to truncate the crustal fabric. Regardless of how this fabric may have originated, the geometry of its structural discordance with respect to the Moho (analogous to an inverted unconformity beneath blocks VII and IX) requires that the Moho, as seen today, formed subsequently.

The above interpretation implies that, at the time of its formation, the Moho constituted a thermo-mechanical boundary between rigid crust and plastic mantle. Any relief on the Moho, such as may have been induced by folding or faulting and tilting of crust, was subsequently erased by ductile flow, whereas the structure was preserved in the cooler crust above.

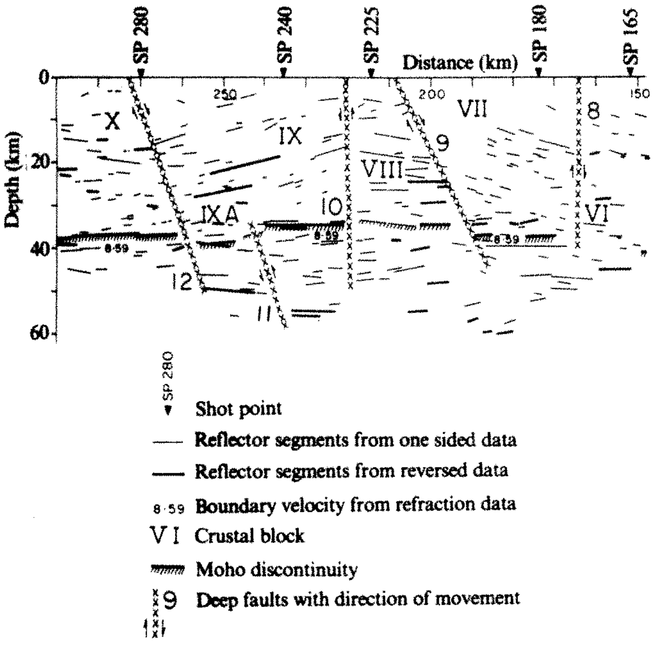

SP 280
SP 240
SP 225
SP 180
SP 165
Distance (km)
Depth (km)
X
IX
IXA
VIII
VII
VI
8·59
Shot point
Reflector segments from one sided data
Reflector segments from reversed data
8·59 Boundary velocity from refraction data
VI Crustal block
Moho discontinuity
9 Deep faults with direction of movement

Fig. 3 Detailed crustal structure for the east-central part of the Kavali–Udipi profile (after ref. 5).

Ductile flow requires that the then Moho level temperatures were close to the crustal solidus. For anhydrous granulite facies rocks, which are believed to constitute the lower crust, ductile flow might be expected at temperatures around 1,000 °C (ref. 13). Consequently, mean geothermal gradients must have been in the range of 20–30 °C km^{-1}, two to three times those inferred for the present day shields[14]. Note that a very similar thermal regime has been inferred from considerations of phase-equillibria in the Lewisian terrain both during the major metamorphism and later at the time of intrusion of Scourie dykes[13]. It is also within the bounds determined from the analysis of *P–T* data for many Archaean high-grade terrains[15].

If the widespread horizontality of the Moho is the result of ductile smearing, then the Moho offsets that are visible today must have arisen later, and their preservation implies that the lithosphere had become cooler by then, with flow possible only at greater depths. This implies that at the time of smearing the Moho was coincident with the base of the lithosphere.

Lithosphere could be described in two ways, depending on the criterion used: 'thermal' lithosphere may be defined as the zone in which heat transfer takes place predominantly by conduction, whereas the thickness of the 'elastic' lithosphere is a function of the change of horizontal shear stress with depth[16]. Starting from a fluid sphere, where they have zero thickness, both grow at different rates with progressive cooling and solidification. Although the elastic lithosphere is much thinner beneath the continents today than the thermal lithosphere, in the early Precambrian they could not have differed much. The smearing-out of relief by ductile flow, postulated above, would, strictly speaking, reflect the base of the elastic lithosphere.

Support for our hypothesis is provided by DSS investigations elsewhere. Central and southeastern Europe have been extensively covered by this technique[1]. In addition to the traverses across Alpine, Hercynian and Caledonian zones, the areas investigated include the Ukrainian Shield (ref. 1, p 19–23). The latter has been so densely covered by DSS profiles, that a Moho contour map could be compiled (ref. 1, Fig. 25). Figure 4 shows the deep crustal section along part of the profile IP I which extends roughly north–south (right to left) across the Ukrainian Shield. The northern part traverses the exposed part of the Shield, beneath which the Moho is seen to be essentially horizontal, and intra-crustal dips are again suggestive of a structural discordance. To the south, the Moho is seen to be dipping. This area is, however, covered by sediments which thicken towards the metasedimentary complex beneath the Scythian Platform. This complex has been multiply deformed in late Proterozoic and Phanerozoic times (ref. 1, pp. 61, 62). We

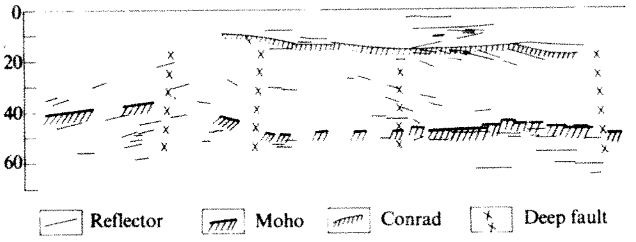


Fig. 4 Simplified crustal section along a part of the profile IP I in the Ukrainian Shield (after ref. 1).

consider that the Moho dips on the left half of Fig. 4 herald the approach to this orogenic zone where deformation occurred when the lithosphere was cooler and thicker, and thus folding of the Moho was preserved.

Preliminary analysis of the near vertical range COCORP data from North Texas, Wyoming and New Mexico in North America has also yielded evidence of a gently dipping Moho[17]. The occurrence of this predominantly horizontal structural style has been ascribed to the Moho under the continents being "a buoyancy driven sharp density contrast (re-) established in a region undergoing strong horizontal shear"[11]. This interpretation is consistent with our hypothesis.

We thank Rishi Narain Singh and Edward S. Grew for helpful discussions. Shankeraiah, Jamaluddin and Rana drew the figures. R.B.H. acknowledges the support of an Indo–US Subcommission fellowship. K.R.C. thanks the director, NGRI for permission to publish.

Received 16 December 1980; accepted 16 April 1981.

1. Sollogub, V. B., Prosen, D. & Militzer, H. (eds) *Geophys. Trans. Hungarian Geophys. Inst.* (spec. edn 1972).
2. Bamford, D. & Prodehl, C. *J. geol. Soc.* **134**, 139–151 (1977).
3. Condie, K. C. *Plate Tectonics and Crustal Evolution* (Pergamon, New York, 1976).
4. Fyfe, W. S. *Nature* **249**, 338 (1974).
5. Kaila, K. L. *et al. J. geol. Soc. Ind.* **20**, 307–333 (1979).
6. Crawford, A. R. *J. geol. Soc. Ind.* **10**, 117–166 (1969).
7. Venkatasubramaniam, V. S. & Narayana Swami, R. *J. geol. Soc. Ind.* **15**, 77–80, 200–204, 318–319 (1974).
8. Ikramuddin, M. & Stueber, A. M. *Lithos* **9**, 235–240 (1976).
9. Beckinsale, R. D., Drury, S. A. & Holt, R. W. *Nature* **283**, 469–470 (1980).
10. Crawford, A. R. & Compston, W. *J. geol. Soc. Austr.* **19**, 453–464 (1973).
11. Phinney, R. A. *EOS* **59**, 389 (1978).
12. Smithson, S. B. *Contr. Geol. Univ. Wyoming* **17**, 65–75 (1979).
13. O'Hara, M. J. *J. geol. Soc.* **134**, 185–200 (1977).
14. Rao, R. U. M., Rao, G. V. & Narain, H. *Earth planet. Sci. Lett.* **30**, 57–64 (1976).
15. Watson, J. V. *Phil. Trans. R. Soc.* **A288**, 431–440 (1978).
16. Phillips, R. J. & Ivins, E. R. *Phys. Earth planet. Inter.* **19**, 107–148 (1979).
17. Cook, F. A., Brown, L. D. & Kaufmann, S. *EOS* **59**, 389 (1978).
18. Naqvi, S. M., Rao, V. D. & Narain, H. *Precambr. Res.* **6**, 323–345 (1978).

The earliest known Palaeocene mammal fauna and its implications for the Cretaceous–Tertiary transition

J. David Archibald

Department of Biology and Peabody Museum of Natural History,
Yale University, New Haven, Connecticut 06520, USA

Several localities in northeastern Montana, USA, representing the earliest known Palaeocene mammal sites in the world, document a gradual turnover in the mammalian fauna at the Cretaceous–Tertiary boundary. These localities, in combination with previously known latest Cretaceous localities, represent the only well documented terrestrial sequence across the Cretaceous–Tertiary boundary. This sequence argues for gradual terrestrial extinctions and faunal changes at this boundary.

The extensively studied latest Cretaceous Lance local fauna, Lance Formation, Wyoming, USA[1], and Trochu local fauna, Scollard Formation, Alberta, Canada[2], are dominated by ptilodontoid multituberculates and marsupials. Proteutherian eutherians are somewhat more taxonomically diverse at the Alberta sites than at those in Wyoming, but at neither are eutherians the dominant mammals. These mammalian faunas are considered typically latest Cretaceous. Their composition differs markedly from early Palaeocene (Puercan) faunas.

Localities which yield only typical latest Cretaceous mammals are also known from McCone County, Montana, USA and have been even better sampled from localities immediately to the west in Garfield County[3]. These localities belong to the Hell Creek faunal-facies which is found throughout most of the vertical and lateral extent of the Hell Creek Formation and which includes sediments deposited in small streams, flood plains and back swamps[3,4] (Fig. 1).

In marked contrast, a series of three biostratigraphically positioned latest Cretaceous localities (Bug Creek Anthills, Bug Creek West, Harbicht Hill) in McCone County, Montana (Fig. 1) have yielded mammals of a distinctly Palaeocene aspect, including condylarths, taeniolabidoid multituberculates, a proteutherian ancestral to taeniodonts, and a primate[5]. These localities belong to the Bug Creek faunal-facies which is found only in the upper 24 m of the Hell Creek Formation and which includes sediments deposited in large streams[3,4] (Fig. 1). Typical latest Cretaceous mammals and dinosaurs occur in this faunal-facies but they become taxonomically less diverse and numerically less abundant up-section as the Palaeocene-aspect mammals increase in both aspects[3,6].

Table 1 Hell's Hollow local fauna, Garfield County, northeastern Montana, USA

- Allotheria—Multituberculata
  - Ptilodontoidea
    - Neoplagiaulacidae
      - *Mesodma*, n. sp.
  - Taeniolabidoidea
    - Eucosmodontidae
      - *Stygimus* aff. *Stygimys kuszmauli*
      - Microcosmodontinae, n. gen. et n. sp.
    - Taeniolabididae
      - *Catopsalis* n. sp.
  - Suborder and family undetermined
    - *Cimexomys minor*
    - *Cimexomys* n. sp.
- Theria
  - Metatheria—Marsupialia
    - Didelphidae
      - *Peradectes* cf. *Peradectes pusillus*
  - Eutheria—order undetermined (proteutherian)
    - Palaeoryctidae
      - *Procerberus formicarum*
  - Eutheria—Condylarthra
    - Arctocyonidae
      - *Protungulatum* cf. *Protungulatum donnae*
      - *Ragnarok harbichti*?
      - *Ragnarok* n. sp.
    - ?Hyopsodontidae
      - *Oxyprimus erikseni*
    - Periptychidae
      - *Mimatuta morgoth*

The major problem in studying the terrestrial vertebrate transition at the Cretaceous–Tertiary boundary has been the lack of closely spaced localities on both sides of the boundary within a geographically isolated region. The same is true for shallow-water marine macroinvertebrates in North America (K. Waage, personal communication). Early Palaeocene (Puercan) mammal localities were previously discovered in both McCone[5] and Garfield[7] counties, but 50 m of section separate these from the latest Cretaceous localities. The discovery of new basal Palaeocene localities in Garfield County separated by only 14 m from latest Cretaceous localities now provides the most closely sampled series of terrestrial vertebrate faunas at the Cretaceous–Tertiary boundary in the world.

The two most fossiliferous of the new basal Palaeocene localities occur in a channel deposit ~5 m above the Cretaceous–Tertiary boundary (as defined by the highest occurrence of dinosaurs) and together comprise the Hell's Hollow local fauna. This local fauna contains 12 or possibly 13 mammalian species and an undetermined number of lower vertebrates. Based on biostratigraphical, and to a lesser extent stratigraphical evidence, this local fauna is older than other early

0028-0836/81/260650—03$01.00

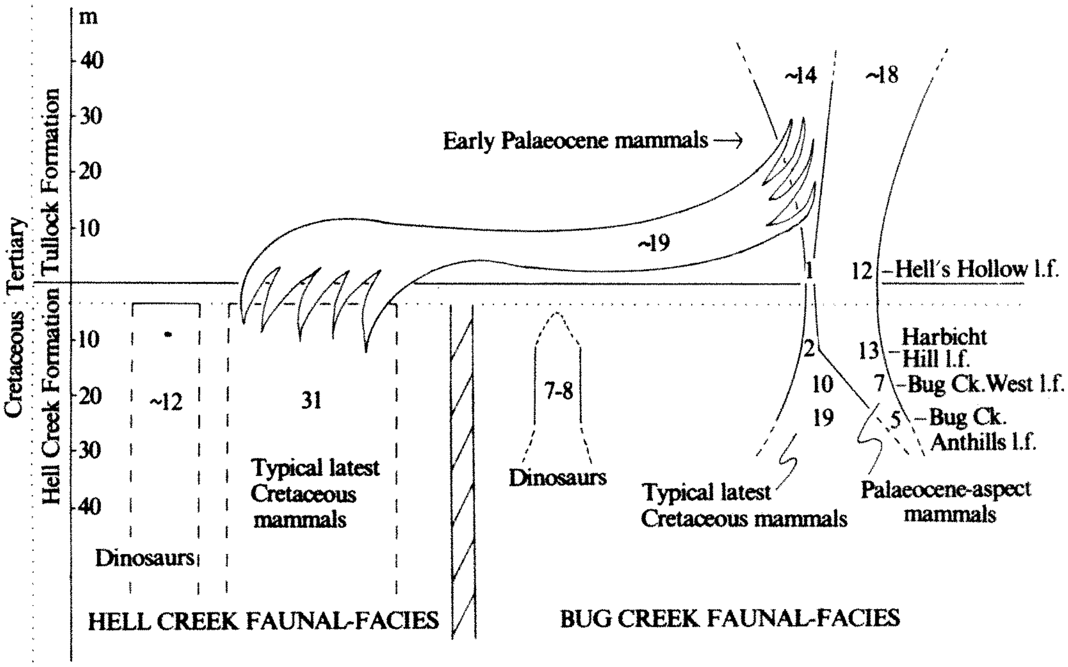


Fig. 1 Stratigraphical and biostratigraphical sequence in northeastern Montana showing relationships of the Hell's Hollow l.f. (local fauna) to the Bug Creek faunal-facies sequence—Bug Creek Anthills l.f., Bug Creek West l.f., Harbicht Hill l.f. (ref. 5), the Hell Creek faunal-facies and the later early Palaeocene sites. Numbers and widths of bars approximate numbers of species known at various levels. Numbers for Bug Creek faunal-facies dinosaurs are from ref. 6, numbers for the highest early Palaeocene mammals from refs 5 and 9, and the remainder from ref. 3.

Palaeocene faunas known from western North America. Faunally it is most similar to the local fauna which is stratigraphically highest in the Bug Creek faunal-facies—the Harbicht Hill local fauna, which is latest Cretaceous in age (Fig. 1).

Unlike Harbicht Hill, Hell's Hollow lacks dinosaurs. Of the two typical latest Cretaceous mammal species remaining at Harbicht Hill, the multituberculates *Meniscoessus robustus* and *Cimexomys minor*, only the latter species is found in Hell's Hollow. Species of the multituberculates *Mesodma* and *Cimexomys*, the marsupial *Peradectes*, and possibly the proteutherian *Procerberus* that occur in Hell's Hollow may be derived from known typical latest Cretaceous mammals. All except *Peradectes* are represented by the same or closely related species at Harbicht Hill. Hell's Hollow records the earliest definite occurrence of this typically Palaeocene marsupial. The remaining Palaeocene aspect mammals (Table 1) or their ancestors first appear as immigrants in the lowest Bug Creek faunal-facies local fauna (Bug Creek Anthills). Except for the Asian origin of the multituberculate *Catopsalis*[8], little is known about the origin of these remaining taxa. The species of the multituberculate *Stygimys* and of the condylarths[9] *Protungulatum*, *Oxyprimus*, *Mimatuta* and questionably *Ragnarok* that are present at Harbicht Hill seem to continue on into Hell's Hollow. *Ragnarok* is also represented by a new species in Hell's Hollow, as is *Catopsalis*. This same species of *Catopsalis* is found at slightly younger localities in Colorado and Wyoming (M. Middleton, personal communication). A new genus and species of microcosmodontine from Hell's Hollow represents the earliest record of this subfamily of eucosmodontid multituberculates.

This earliest known Palaeocene local fauna is clearly a continuation of the latest Cretaceous Bug Creek faunal-facies. In turn, the Hell's Hollow local fauna or similar local faunas gave rise to later faunas of the early Palaeocene in western North America. Although there was taxonomic turnover and possibly a decrease in diversity (as suggested in Fig. 1) during the Cretaceous–Tertiary transition in Montana, the changes were gradual. Although the evidence is not as clear, the Hell Creek faunal-facies may have disappeared rather abruptly as a unit at the close of the Cretaceous. However, about 19 of the 31 species known in this faunal-facies or their close relatives occur at Palaeocene localities. Only the marsupials were decimated, dropping from 13 to 1 species. Such selective extinctions point to relatively normal ecological changes, for catastrophic events would not have been so selective.

These new data, combined with previous information on the changes in the terrestrial biota at the Cretaceous–Tertiary boundary,[6] all indicate that this was a non-catastrophic transition. Megaflora, and particularly palynoflora, have been sampled across the Cretaceous–Tertiary boundary at several localities in North America and elsewhere[10]. The only consistent pattern of major change that has been recorded is the decimation of the pollen and megaflora within the *Aquilapollenites* Province from western North America and Siberia at the close of the Cretaceous, although a few species belonging to this group continue into the Palaeocene[10]. Further, in Garfield County, the palynofloral extinctions occurred ~9 m above the extinction of dinosaurs. Similar incongruencies found in Alberta[11], southeastern Wyoming[12] and northern Wyoming (personal communication from P. Gingerich to L. Hickey) also point to a recognizably slow transition of the terrestrial biota at the end of the Cretaceous.

This evidence is at least in part compatible with hypotheses—climatic change, competition with invading mammals of Palaeocene aspect[6], drops in speciation rates due to habitat redistribution[13]—that argue for a gradual transition. Although it is difficult to estimate the duration of this transition, those that have been suggested place the magnitude in the range of tens or hundreds of thousands of years[9,11], not in the tens or hundreds of years suggested by various catastrophic hypotheses[14,15]. Several of these recent hypotheses have presented data which suggest that an extraterrestrial body struck or closely approached the Earth at the close of the Cretaceous and that its biological effects were catastrophic[14–16]. The chemical and physical evidence for such an event is intriguing and worthy of further pursuit, but the palaeontological evidence cannot be used to evaluate these hypotheses directly. However, it is only through palaeontological investigation that the biological events at the Cretaceous–Tertiary boundary can be investigated.

Comments and data provided by W. A. Clemens, L. Hickey, J. Ostrom, K. Padian, S. Rachootin, D. Schindel and K. Waage are gratefully acknowledged. Much of the field work was supported by NSF grant DEB 75-21027 to W. A. Clemens.

Received 19 January; accepted 24 April 1981.

1. Clemens, W. A. *Univ. Calif. Publs geol. Sci.* **48,** 1–105 (1964); **62,** 1–122 (1966); **94,** 1–102 (1973).
2. Lillegraven, J. A. *Paleont. Contr. Univ. Kans.* **50** (Vert. 12), 1–122 (1969).
3. Archibald, J. D. *Univ. Calif. Publs geol. Sci.* (in the press).
4. Clemens, W. A. & Archibald, J. D. *Mém. Soc. géol. Fr.* **139,** 67–74 (1980); Archibald, J. D. & Clemens, W. A. in *The New Uniformitarianism: Events on a Geologic Scale* (eds Berggren, W. A. & Van. Couvering, J.) (Princeton Press, in the press).
5. Sloan, R. E. & Van Valen, L. *Science* **148,** 220–227 (1965); Van Valen, L. & Sloan, R. E. **150,** 743–745 (1965).
6. Van Valen, L. & Sloan, R. E. *Evolut. Theory* **2,** 37–64 (1977).
7. Clemens, W. A. *Science* **184,** 903–905 (1974).
8. Kielan-Jaworowska, Z. & Sloan, R. E. *Acta palaeont. pol.* **25,** 187–197 (1979).
9. Van Valen, L. *Evolut. Theory* **4,** 45–80 (1978).

10. Hickey, L. in *The New Uniformitarianism: Events on a Geologic Scale* (eds Berggren, W. A. & Van Couvering, J.) (Princeton Press, in the press).
11. Lerbekmo, J. F., Evans, M. E. & Baadsgaard, H. *Nature* **279**, 26–30 (1979).
12. Leffingwell, H. A. *Geol. Soc. Am. Spec. Pap.* **127**, 1–64 (1970).
13. Bakker, R. T. in *Patterns of Evolution, as Illustrated by the Fossil Record* Vol. 5 (ed. Hallam, A.) 439–468 (Elsevier, Amsterdam, 1977).
14. Alvarez, L. W., Alvarez, W., Asaro, F. & Michel, H. V. *Science* **208**, 1095–1108 (1980).
15. Smit, J. & Hertogen, J. *Nature* **285**, 198–200 (1980).
16. Hsü, K. J. *Nature* **285**, 201–203 (1980).

Paternity and inheritance of wealth

John Hartung*

Department of Anthropology, Harvard University, Cambridge,
Massachusetts 02138, USA

One of the oldest conjectures in anthropology is that men transfer wealth to their sister's son when the biological paternity of their 'own' children is in doubt[1–12]. Because maternity is certain, a man is necessarily related to his sister's son and his brother (see Fig. 1). It is argued here that relatedness to male heirs can be assured by passing wealth to sister's sons or down a line of brothers, whether the prevailing kinship system reckons those brothers matrilineally or patrilineally. It is also argued that when several transfers of wealth are considered, a man's likelihood of being cuckolded need not be unrealistically high[13] for his successive matrilineal heirs to be more related to him than his successive patrilineal heirs (see Fig. 2). Cross-cultural data on sister's son/brother inheritance[14] and frequency of extramarital sex for females[15] support the hypothesis that men tend to transmit wealth to their sister's son and/or brother when the probability that their putative children are their genetic children is relatively low.

In matrilineal inheritance a man's primary heir is his sister's son. This cross-generation transfer can be direct (25 societies) but more often occurs after wealth has been sequentially inherited by younger brothers ($\simeq 47$ societies), the last of whom leaves it to their mutual sister's son. Although matrilineal inheritance does not prevent a man from distributing significant benefits to his own children[16,17], and is thereby not a mirror image of patrilineal inheritance[18] ($\simeq 500$ societies), it is unique in designating major cross-generation heirs related to a male benefactor through an unbroken line of females (counts are of a composite inheritance score derived from ref. 14, as described below).

In 1771, after characterizing traditional societies as patrilineal with regard to inheritance of wealth, status and family membership, John Millar[1] noted an exception: "We read of several nations, among whom it [marriage] is either unknown, or takes place in a very imperfect and limited manner. To a people in this situation it will appear that children have much more connection with their mother than with their father. If a woman has no notion of attachment and fidelity to any particular person, if notwithstanding her occasional intercourse with different individuals . . . the child which she has born is regarded as a member of her own family . . . so that if any person was desired to give an account of the family to which he belonged, he was naturally led to recount his maternal genealogy in the female line. [In such nations] females, indeed, are held incapable of enjoying the office of chief, but through them the succession to that dignity is continued; and therefore, upon the death of a chief, he is succeeded not by his own son, but by that of his sister." In 1865 McLennan[3] renewed the hypothesis with enthusiasm (indeed, over-enthusiasm): "The connection between these two things—uncertain paternity and kinship through females only, seems so necessary—that of cause and effect—that we may confidently infer the one where we find the other."

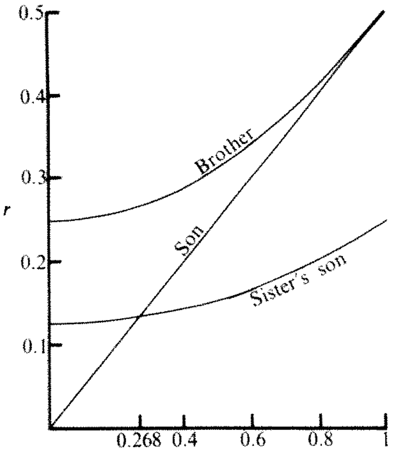


Fig. 1 p = Probability of paternity = the probability that a man's putative offspring by his wife are, in fact, his biological offspring; r = probabilistic degree of relatedness. If p is below 0.268, a man is more related to his sister's son than his 'own' son—see ref. 9 and Fig. 2. A man is more related to his brother whenever p is <1. r to son $=\frac{1}{2}p$; r to sister's son equals r through mother to sister to sister's son $=(\frac{1}{2})(\frac{1}{2})(\frac{1}{2})$, plus r through father to sister to sister's son $=(\frac{1}{2}p)(\frac{1}{2}p)(\frac{1}{2})$, which together $=(\frac{1}{2})(\frac{1}{4})(1+p^2)$; r to brother equals r through mother $=\frac{1}{4}$, plus r through father $=\frac{1}{4}p^2$, which together $=\frac{1}{4}+\frac{1}{4}p^2$.

In 1961 the first empirical study of matrilineality[19] found that it is distributed independently around the world and is associated with fishing, while patrilineality is associated with pastoralism—an intriguing finding in light of another of Millar's observations[1]: "Among the inhabitants of the eastern coast of Tartary it is said, that those tribes which are accustomed to the pasturing of cattle discover some sort of jealousy with regard to the chastity of their women, a circumstance which is looked upon as a matter of perfect indifference by those families in that country who procure their subsistence merely by fishing."

In 1974 Alexander[7] made the hypothesis explicit with regard to modern evolutionary theory, and in 1978 Greene[9] calculated the probability of paternity (p) below which husbands must fall if they are to be, on average, more related to their sister's sons than their 'own' sons: $p < 0{\cdot}268$ (Fig. 1). This threshold is so low because a man's probable relatedness to his sister's son must be reduced by his probable relatedness to his sister through their father, that is, the probability that he is his father's son and she his father's daughter. Unfortunately (for the hypothesis), if a society-wide p were < 0.268, the average man would have more than twice as much reproductive success by other men's wives as he would by his own wife—a circumstance which, in addition to straining the imagination, could render marriage a maladaptive male strategy (probably better to remain single and expend all of one's efforts on other men's wives).

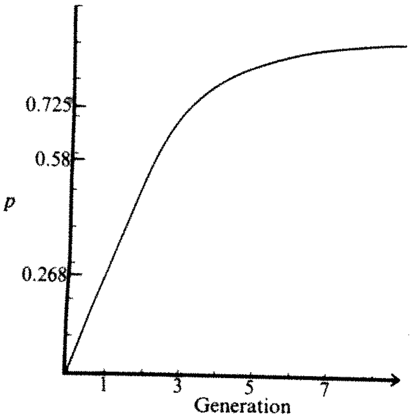
Fortunately, a less extreme p can fit the paternity hypothesis if more than one generation is considered when contrasting ideal matrilineality against ideal patrilineality (father to son, to son's son, to son's son's son, etc.). The key here is that while a man with a p of 0.8 is probabilistically related to his son by $0.8 \times r$ where r is the degree of relatedness, he is related to his son's son by $0.8^2 \times r$ and to his son's son's son by $0.8^3 \times r$ (where $r = 0.5$, 0.5^2 and 0.5^3, respectively). That is, the exponent of p increases by a factor of 1 for each transfer to an heir related through a male. In contrast, p affects only the first cross-generation matrilineal transfer. A man's sister's son's sister's son is related to him by half as much as his sister's son, with no additional effect from p. This means, for example, that if p is < 0.725, a man is more related to his matrilineal heir than his patrilineal heir at three generations (Fig. 2).

Assuming that the paternity hypothesis is viable for reasonable values of p, assessing it would seem a matter of cross-tabulating matrilineal inheritance versus patrilineal inheritance by high versus low extramarital sex for females. However, there is a form of 'delayed' patrilineal inheritance, patrilineal brother inheritance, which can be seen as an alternative adaptation to low p. Given any $p < 1$, a man is more related to his brother than to any other potential male heir (Fig. 1). Because p cannot be 1, one might expect all patrilineal benefactors to follow the pattern

---

* Present address: 2G, 160 Claremont Ave, New York, New York 10027, USA.

0028–0836/81/260652—03$01.00



Generation	r to sister's son$_i$	r to son$_i$	Threshold at $p=$
1	$(\frac{1}{2})(\frac{1}{4})(1+p^2)$	$\frac{p}{2}$	0.268
2	$(\frac{1}{2})^2(\frac{1}{4})(1+p^2)$	$\left(\frac{p}{2}\right)^2$	0.58
3	$(\frac{1}{2})^3(\frac{1}{4})(1+p^2)$	$\left(\frac{p}{2}\right)^3$	0.725
i	$(\frac{1}{2})^i(\frac{1}{4})(1+p^2)$	$\left(\frac{p}{2}\right)^i$	$4p^i-p^2-1=0$

Fig. 2 Solving for p given i, $4p^i-p^2-1=0$, where $i=i$th generation, derives the probability of paternity below which a man is more related to his cross-generation matrilineal heir (sister's son$_i$) than his cross-generation patrilineal heir (son$_i$) at successive generations. For example, if p is <0.725, a man is likely to be more related to his third generation matrilineal heir than to his son's son's son. (I thank P. Greene for deriving this formula.)

of giving to their younger brother(s) with the stipulation that the wealth subsequently return to their son(s). However, at an average father's age of death, his son is near his reproductive prime while his brother is past that stage. Accordingly, a son's higher reproductive value[20] will compensate, in terms of effect on benefactor's inclusive fitness[21], when p is <1 but not $\ll 1$.

Why adaptation to low p should be patrilineally reckoned brother inheritance in some societies ($\simeq 70$ cases) and matrilineal inheritance in other societies ($\simeq 72$ cases) is unclear. The absolute relatedness of a brother is higher than that of a sister's son, but in patrilineal brother inheritance wealth eventually reverts to a man's own child. There seems to be a trade-off between absolute relatedness to heir and the number of transfers for which relatedness is assured (with matrilineal inheritance assuring some relatedness indefinitely). Perhaps the two strategies reflect a difference in p, or economic differences which affect the duration of effectiveness of inherited wealth, or perhaps patrilineal brother inheritance is simply as 'matrilineal' as an otherwise patrilineal society can get. Whatever the reason, given brother inheritance, whether in a matrilineal or patrilineal context, any inclusive fitness enhancing effect of wealth[8,22–24] resides for the longest possible time amongst most highly related males (for emphasis on males see refs 8, 22, 25, 26). The key here is that, for present purposes, patrilineal brother inheritance has more in common with matrilineal inheritance than with patrilineal inheritance. Thus, one should group patrilineal brother inheritance with matrilineal inheritance when testing the hypothesis that men stipulate an heir whose relatedness is less affected by low p in societies where p is relatively low.

Murdock and White have selected 'The Standard Cross-Cultural Sample'[27] by taking one society from each of the world's 186 major culture areas, giving a sample which maximizes independence and minimizes effects from cultural diffusion. Separate codes for inheritance of real property (land, house) and movable property (such as livestock, money) are as follows[14,28,29]:

0: No information.
1: No land rights/no inherited movable property/no rule governing the transmission of same.
2: Matrilineal inheritance by a sister's son or sons.
3: Matrilineal inheritance by heirs who take precedence over sisters' sons (such as younger brother).
4: Inheritance by children, but with daughters receiving less than sons.
5: Inheritance by children of either sex or both.
6: Inheritance by patrilineal heirs who take precedence over sons (such as younger brother).
7: Patrilineal inheritance by son or sons.

For present purposes code 3 is functionally the same as code 2—though brother is not stipulated as the only possibility, he is the most likely heir[16,18]. Where this is not the case, the pattern still qualifies as having heirs unaffected by low p, because any matrilineal heir must be related through the mother and because

PRIMARY HEIR(S)

		Brother/sister's son		Own children
FEMALE EXTRAMARITAL SEX	high	A 6 Nyakyusa Nkundo Lesu Trobrianders Saramacca Crow	Huichol?	B 5 Romans Vietnamese Kwoma Trukese Aymara
	low	C 1 Siuai		D 10 Kikuyu Egyptians Turks Irish Lapps Lepcha Lakher Vedda Javanese Iban

Fig. 3 Inheritance by probability of paternity. $N=22$; Fisher exact probabililty $=0.0298$; $\gamma=0.846$ (considering the two inheritance forms as 'high' and 'low' with regard to effect of p on relatedness to heir); $\phi=0.488$. The Huichol (see text) are rated 'high' on female extramarital sex and would go in cell A or B if included—either way the figure would retain significance below 0.05. This significance depends on grouping patrilineal brother inheritance with matrilineal inheritance. If the contrast were matrilineal versus own children/patrilineal brother, the Nyakyusa, Nkundo and Huichol would fall in cell B; giving cell $A=4$, cell $B=8$, cell $C=1$, cell $D=10$: Fisher exact probability $=0.162$. If the contrast were simply matrilineal versus own children, the Nyakyusa, Nkundo and Huichol would drop out, giving cell $A=4$, cell $B=5$, cell $C=1$, cell $D=10$: Fisher exact probability $=0.0893$. Murdock expressed doubt about his inheritance codes[14]. In response to an inquiry he wrote that he did not have reservations about the accuracy of the data, but felt that the codes were 'too crude' (personal communication). Future research would profit from refining the inheritance codes. For example, a man could extend considerable assistance to his brother, both during his lifetime and through inheritance and still follow the norms of a society coded in inheritance category 'own children'. According to the paternity hypothesis as presented here, societies in cell B would be expected to follow that pattern more than societies in cell D, and some cell B societies may have sufficient 'brother investment' to qualify for cell A.

wealth eventually goes to sister's son. Societies with code 6, however, must be checked to see whether brother is the heir. Of the 23 cases with usable codes for inheritance (see below) and frequency of female extramarital sex, three contain a 6 in their composite inheritance score: the Nyakyusa (6 and 6), the Nkundo (6 and 6) and the Huichol (5 and 6). These were checked in the Human Relations Area Files[28] to see whether brother is the preferred heir. The major ethnographer for the Nyakyusa puts the case succinctly[30]: "A man's heir is his full brother next to him in age, and only when all the full brothers of a family have died does the property pass to the eldest son of the eldest brother." For the Nkundo, cases were reported of brother taking precedence over son, of sister's son taking precedence over son, and of brother taking precedence over sister's son[31]. As put by Murdock[32], the Nkundo "lie wedged between the matrilineal Central Bantu in the south and the patrilineal Equatorial Bantu in the north. Being unusually well described, they shed considerable illumination on the process whereby the transition from the one to the other form of social organization has occurred in this general region." Perhaps the Nkundo would have to be coded 'none of the above', but it is clear that their system provides substantial brother, if not matrilineal, inheritance. The case for the Huichol is not as clear. Movable property goes to children, but real property goes communally to a group of family members that probably, but not necessarily, contains brother(s)[33]. Accordingly, the Huicol were left out of the analysis.

To maximize sample size and variable dichotomy, the contrast for inheritance norms was made between: (1) societies that stipulate heirs necessarily related to male benefactors through females (heirs whose r to benefactor is relatively unaffected by low p), and (2) societies that exclude such heirs and stipulate only those heirs whose r to benefactor is highly affected by low p. Accordingly, societies were grouped into: (1) those with brother/matrilineal inheritance for either type of property (code 2, 3 or 6, where 6 has been checked—see above), or (2) those for which benefactor's own children are the only stipulated heirs for both types of property (code 4, 5 or 7).

Broude and Greene[15] have coded frequency of extramarital sex for females as: (1) universal, moderate, and (2) rare, not common[34]. As would be predicted by the paternity hypothesis, the appropriate cross-tabulation (Fig. 3) indicates an association between relatively frequent female extramarital sex and a cultural norm that allows men to designate heirs whose relatedness is relatively unaffected by low probability of paternity.

It is assumed that the behaviours relevant to this analysis have little heritability across same-sex individuals within a society, and negligible heritability by sex across societies—that is, it is presumed that differences in individual behaviour are phenotypic responses to different conditions, that the behaviours in question have been too important for too long for there to be much genetic variance left, and that extant within-sex genetic variance is an irrelevant variable (compare refs 26, 35 with 36).

I thank D. Aberle, R. D. Alexander, N. Chagnon, I. DeVore, M. Dickeman, P. Ellison, P. Greene, W. Irons, P. Harvey, G. P. Murdock, J. Shepher, B. Smuts, P. van den Berghe and J. Whiting for their thoughts and assistance. Particular thanks are due to N. Schmidt for her impetus in re-acquiring the HRAF at Harvard.

Received 2 September 1980; accepted 8 April 1981.

1. Millar, J. *Observations Concerning the Distinction of Ranks in Society*, 30, 33, 42 (Murray, London, 1771).
2. Bachoffen, J. J. *Das Mutterrecht* (Beno Schwabe, Basel, 1861).
3. McLennan, J. F. *Primitive Marriage*, 161 (Black, Edinburgh, 1865).
4. Morgan, L. H. *Ancient Society* (Holt, New York, 1877).
5. Junod, H. *Life of a South African Tribe* (Attinger Feres, Neuchatel, 1913).
6. Kahn, M. C. *Djuka: the Bush Negroes of Dutch Guiana* (Viking, New York, 1931).
7. Alexander, R. D. *A. Rev. Ecol. Syst.* **5**, 324–383 (1974).
8. Hartung, J. *Curr. Anthrop.* **17**, 4, 607–622 (1976).
9. Greene, P. J. *Am. Ethnol.* **5**, 1, 151–159 (1978).
10. Kurland, J. A. in *Evolutionary Biology and Human Social Behavior* (eds Chagnon, N. A. & Irons, W.) (Duxbury, North Scituate, Massachusetts, 1979).
11. Flinn, M. in *Natural Selection and Social Behavior: Recent Research and New Theory* (eds Alexander, R. D. & Tinkle, D. W.) (Chiron, New York, 1981).
12. Gaulin, J. C. & Schlegel, A. *Ethol. Sociobiol.* **1**, 301–309 (1980).
13. Swetnam, J. J. & Knack, M. C. *Am. Anthrop.* **81**, 108 (1979).
14. Murdock, G. P. *The Ethnographic Atlas* 1,170 case punch card version with codebook, HRAF (University of Pittsburgh Press, 1967).
15. Broude, G. J. & Greene, S. J. *Ethnology* **15**, 409–429 (1976).
16. Richards, I. A. in *African Systems of Kinship and Marriage* (eds. Radcliffe-Brown, A. R. & Forde, D.) 207–251 (Oxford University Press, London, 1950).
17. Panoff, M. *Ethnology* **15**, 175–188 (1976).
18. Schneider, D. & Gough, K. (eds) *Matrilineal Kinship* (University of California Press, 1961).
19. Aberle, D. in *Matrilineal Kinship* (eds. Schneider, D. & Gough, K.) 655–727 (University of California Press, 1961).
20. Fisher, R. A. *The Genetical Theory of Natural Selection* (Oxford University Press, 1930).
21. Hamilton, W. D. *J. theor. Biol.* **7**, 1–52 (1964).
22. Irons, W. in *Evolutionary Biology and Human Social Behavior* (eds Chagnon, N.A. & Irons, W.) (Duxbury, North Scituate, Massachussetts, 1979).
23. Keyfitz, N. *Applied Mathematical Demography* (Wiley, New York, 1979).
24. Frisch, R. *Science* **199**, 22–30 (1978).
25. Hartung, J. *Curr. Anthrop.* (in the press).
26. Hartung, J. in *Natural Selection and Social Behavior: Recent Research and New Theory* (eds Alexander, R. D. & Tinkle, D. W.) (Chiron, New York, 1981).
27. Murdock, G. P. & White, D. R. *Ethnology* **8**, 329–369 (1969).
28. HRAF, Atlas codebook (see ref. 14).
29. Textor, R. B. *A Cross-Cultural Summary*, 112 (HRAF, New Haven, 1966).
30. Wilson, M. *Good Company: a Study of Nyakyusa Age Villages*, 35 (Oxford University Press, London, 1951).
31. Hulstaert, G. E. *Le Mariage des Nkundo* (G. van Campenhout, Bruxelles, 1928).
32. Murdock, G. P. *Africa—Its People and their Culture History*, 284–289 (McGraw-Hill, New York, 1959).
33. Znigg, R. M. *The Huichols: Primitive Artists* (Stechert, New York, 1938).
34. Broude, G. J. *Behav. Sci. Res.* **4**, 227–262 (1976).
35. Hartung, J. *Curr. Anthrop.* **21**, 1, 131 (1980).
36. Lumsden, C. J. & Wilson, E. O. *Genes, Mind and Culture: The Coevolutionary Process* (Belknap/Harvard Press, Cambridge, 1981).

Sustained potentiation of transmitter release by adrenaline and dibutyryl cyclic AMP in sympathetic ganglia

K. Kuba[*], E. Kato[†], E. Kumamoto[*], K. Koketsu[†] & K. Hirai[†]

[*] Department of Physiology, Saga Medical School, Nabeshima, Saga 840-01, Japan
[†] Department of Physiology, Kurume University School of Medicine, Kurume 830, Japan

Cyclic AMP[1] is known to mediate actions of various hormones and neurotransmitters in a variety of cells[2–5]. Its physiological role, however, in the transmitter release mechanism of vertebrate synapses is still controversial[6–9]. We report here evidence for the sustained, acceleratory effects of adrenaline and dibutyryl cyclic AMP on transmitter release in bullfrog sympathetic ganglia. The amplitude and quantal content of the fast excitatory postsynaptic potential (e.p.s.p.) and the frequency of miniature excitatory postsynaptic potential (m.e.p.s.p.) increased markedly for a long time after pretreatment with adrenaline or dibutyryl cyclic AMP. These results suggest a significant modulatory role for adrenaline, presumably linked with a cyclic AMP system, in the transmitter release mechanism, which may serve as a mechanism for neuronal plasticity in the peripheral autonomic nervous system.

The ninth or tenth paravertebral sympathetic ganglia of the bullfrog (*Rana catesbeiana*) were isolated. B-type neurones were studied using a conventional intracellular recording technique[10]. Microelectrodes were filled with 3 M KCl or 1 M K_3-citrate (tip resistance 25–100 MΩ). The fast e.p.s.p. generated by the nicotinic action of acetylcholine (see ref. 11) was recorded in a low Ca^{2+}–high Mg^{2+} solution (Ca^{2+} 0.7–0.9 mM, Mg^{2+} 5.4–6.5 mM). The quantal content of the fast e.p.s.p. was calculated by both the variance and failure methods[12] and the average taken; m.e.p.s.ps were recorded in a high-K^+ (10 mM) solution. The composition of normal Ringer was described previously[10] and experiments were done at 20–24 °C.

Figure 1 shows the effects of adrenaline on the fast e.p.s.p. Perfusion of the sympathetic ganglion with a solution containing adrenaline (100 μM) decreased the amplitude of the fast e.p.s.p. ($65 \pm 7\%$ (s.e.) of the control, $n = 5$) and its quantal content

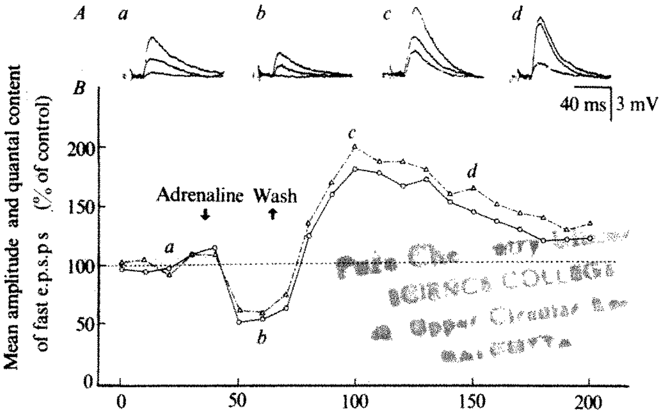


Fig. 1 *A*, effects of adrenaline (100 μM) on the fast e.p.s.ps recorded in a low Ca^{2+}–high Mg^{2+} solution. The amplitude of fast e.p.s.ps in each condition fluctuated in a quantal manner for their low quantal contents. *B*, time course of the adrenaline action. Each point is the mean amplitude (○) or quantal content (△) calculated from 200 fast e.p.s.ps. Quantal contents before application of adrenaline ranged from 1.6 to 2.0. Downward and upward arrows indicate the beginning and end, respectively, of the application of adrenaline. The duration of the 'after-potentiation' of the fast e.p.s.p. by adrenaline observed in this cell was apparently shorter than that induced by dibutyryl cyclic AMP in the other cell (shown in Fig. 2). However, this difference seems to be mainly due to the variability of cell responses, and not to the specificity of the stimulating agents. In fact, a much longer potentiation (~5 h) by adrenaline was seen in other cells (see text). The points *a*, *b*, *c* and *d* in *B* correspond to *a*, *b*, *c* and *d* in *A*, respectively.

(74±7% of the control, $n=5$). Ten to fifteen minutes after washing with an adrenaline-free solution, the amplitude and quantal content of the fast e.p.s.p. recovered and rose to 155% of the control (±8%, $n=5$, the value at 30 min after wash) and to 170% (±14%, $n=5$), respectively. This potentiation of the fast e.p.s.p. continued for >2–5 h (Fig. 1*B*) and was observed with an adrenaline concentration as low as 1 μM. Similar modes of adrenaline action on the fast e.p.s.p. were also observed in Ringer (at normal Ca^{2+} concentration) containing D-tubocurarine (28–43 μM). The quantal size of the fast e.p.s.p., the resting potential and membrane resistance remained almost unchanged throughout the adrenaline actions. These results, together with the lack of effects of adrenaline on acetylcholine-induced depolarization of the ganglion cell membrane, indicate the presynaptic bimodal actions of adrenaline which are quite different from the predominantly depressant effects of adrenaline in mammalian sympathetic ganglia[13]. The delayed onset of the facilitatory effect of adrenaline can be accounted for by the existence of the strong inhibitory action of adrenaline during exposure to it. Thus, without the facilitatory action, the magnitude of the inhibition in the presence of adrenaline would have been much more intense than that observed.

The inhibitory action of adrenaline (10 μM) tended to be antagonized by α-antagonists of the adrenoreceptor, such as phentolamine (10 μM) and phenoxybenzamine (10 μM), but not by a β-antagonist, propranolol (10 μM). On the other hand, the 'after-potentiation' of the fast e.p.s.p. by adrenaline was relatively small or not seen in the presence of either α- or β-antagonist. Furthermore, both dopamine (known to be an α-agonist; 10 μM) and isoprenaline (β-agonist; 10 μM) significantly potentiated the fast e.p.s.p. Thus, it seems premature to determine the type of adrenoreceptor responsible for the facilitation of transmitter release until quantitative and kinetic analyses are made. Alternatively, the receptor for the facilitation mechanism might be a 'primitive' type which cannot be classified into α- or β-type.

The fast e.p.s.p. amplitude tended to be slightly reduced during treatment with dibutyryl cyclic AMP (N^6-$N^{2'}$-*O*-dibutyryl adenosine 3′, 5′-monophosphate; 0.8–1 mM), which is known to be more permeable to the cell membrane than cyclic AMP (Fig. 2 *Ab*). Five to fifteen minutes after the removal of dibutyryl cyclic AMP, the fast e.p.s.p. amplitude increased gradually to 165% of the control (±15%, $n=5$, the value at 30 min after wash). This was accompanied by an increased quantal content (165±13% of the control, $n=5$) without a significant change in the quantal size, resting membrane potential or resistance. Furthermore, 'after-potentiation' was seen at a concentration as low as 0.1 mM. Thus, dibutyryl cyclic AMP has a presynaptic action. As found in the 'after-effect' of adrenaline, the potentiation by dibutyryl cyclic AMP persisted for >2–5 h (Fig. 2*c*, *d*). Similar effects of dibutyryl cyclic AMP were also seen with the fast e.p.s.p. recorded from a curarized ganglion cell in Ringer.

The presynaptic facilitation by dibutyryl cyclic AMP (0.8–1.0 mM) was not observed during treatment for up to 30 min with the drug. By contrast, pretreatment with dibutyryl cyclic AMP even for 5–10 min caused the 'after-potentiation' of transmitter release. Thus, the presence of dibutyryl cyclic AMP at the outer surface of the terminal membrane seems to inhibit the facilitatory mechanism or its initiation. Cyclic AMP, being less permeable to the cell membrane, AMP, ATP and adenosine (4 mM) did not produce 'after-potentiation' of the fast e.p.s.p. Furthermore, 3-isobutyl-1-methylxanthine (IBMX, 20 μM; known to be a phosphodiesterase inhibitor) increased the fast e.p.s.p. amplitude in the ganglion cell which was pretreated with adrenaline (10 μM) 30–40 min before IBMX application.

To test whether the spontaneous release of transmitter is enhanced by adrenaline and dibutyryl cyclic AMP, as with evoked release, the effects of these agents on the m.e.p.s.p. frequency were observed. Although a change in the m.e.p.s.p. frequency in the presence of adrenaline (100 μM) was variable among cells, it was clearly enhanced in many cells after the removal of adrenaline (176±34% of the control at 15 min, $n=8$). Similarly, the m.e.p.s.p. frequency increased after the application of dibutyryl cyclic AMP (1–4 mM) and its subsequent removal (172±27%, $n=9$). These potentiations of spontaneous transmitter release continued for many hours.

Our experimental findings provide evidence of the two hypothetical, but important concepts for the mechanism of vertebrate synaptic transmission. First, the similar modes of presynaptic facilitation by adrenaline and dibutyryl cyclic AMP, but not by less permeable cyclic AMP and other nucleotides, suggest that the action of adrenaline is mediated by endogenous cyclic AMP

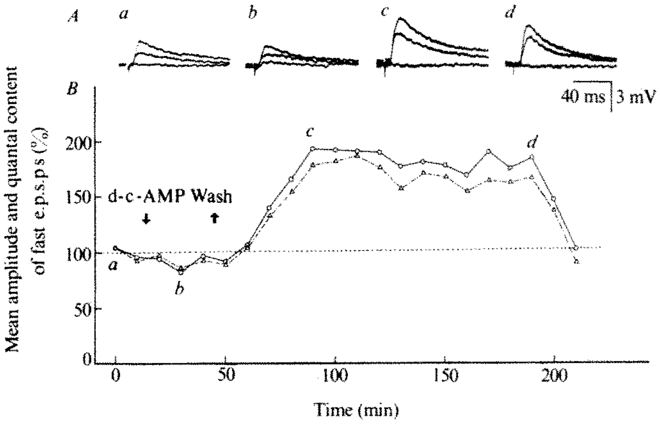


Fig. 2 *A*, effects of dibutyryl cyclic AMP (d-c-AMP) on the fast e.p.s.ps recorded in a low Ca^{2+}–high Mg^{2+} solution. Fluctuations of the fast e.p.s.p. amplitude were seen in each condition for their low quantal content. *B*, time course of the dibutyryl cyclic AMP action. The mean amplitude and quantal content of the fast e.p.s.ps before application of the drug were taken to be 100% in the ordinate. Control quantal contents ranged from 0.7 to 1.0. Each point is the mean of the amplitude (○) or quantal content (△) calculated from 200 fast e.p.s.ps. Downward and upward arrows indicate the beginning and end of the treatment with dibutyryl cyclic AMP (1 mM), respectively.

and a related metabolic process. It has been shown that catecholamines increase cyclic AMP in bovine[14] and rat[15,16] whole sympathetic ganglia, although such an increase in the preganglionic terminals is not yet proved. The latter possibility might be substantiated by the recent observation that dopamine raised the cyclic AMP level in the synaptosome of the guinea pig brain[17]. Thus, adrenaline may raise the cyclic AMP level in the preganglionic terminals, thereby triggering an unidentified mechanism that would eventually modify the Ca^{2+} transport mechanism at the presynaptic terminal membrane[18,19], intracellular Ca^{2+} metabolism[20] or transmitter synthesis[2]. The presynaptic facilitation by serotonin and cyclic AMP in *Aplysia* ganglia was explained by an increase in the active Ca^{2+} conductance during an action potential[18,19]. However, the mechanism of the potentiation seen in the sympathetic ganglia seems to involve at least either an increased level of the intracellular free Ca^{2+} or a raised amount of releasable transmitter, as both the spontaneous and evoked releases of transmitter were enhanced.

The other significant finding is the prolonged maintenance of the presynaptic facilitation, suggesting a plasticity in the peripheral autonomic nervous system. When impulses from the central nervous system activate the postganglionic neurone soma and/or the internuncial adrenergic cells (small intensely fluorescent cells), these would liberate catecholamines. An increased level of catecholamines around the preganglionic terminals might set off the facilitatory mechanism of transmitter release process, and result in the enhanced efficacies of not only the fast (nicotinic), but also the slow (muscarinic) synaptic transmission[11] for many hours. The presynaptic facilitation by adrenaline and dibutyryl cyclic AMP in the bullfrog sympathetic ganglia is quite different from the postsynaptic mechanism for the potentiation of the slow e.p.s.p. by dopamine and cyclic AMP in mammalian ganglia[21,22], possibly due to species difference. As already discussed, however, the slow e.p.s.p. would also be augmented if the transmitter release is potentiated.

Received 3 December 1980; accepted 8 April 1981.

1. Sutherland, E. W. & Rall, T. W. *J. biol. Chem.* **232,** 1065–1076 (1958).
2. Greengard, P. *Nature* **260,** 101–108 (1976).
3. Nathanson, J. A. *Physiol. Rev.* **57,** 157–256 (1977).
4. Rasmussen, H. & Goodman, D. B. P. *Physiol. Rev.* **57,** 421–509 (1977).
5. Phillis, J. W. *J. Can. Sci. Neurobiol.* **4,** 151–195 (1977).
6. Ginsborg, B. L. & Hirst, G. D. S. *J. Physiol., Lond.* **224,** 629–645 (1972).
7. Miyamoto, M. D. & Breckenridge, B. M. *J. gen. Physiol.* **63,** 609–624 (1974).
8. Standaert, F. G., Dretchen, K. L., Skirbol, L. R. & Morgengroth, V. H. III. *J. Pharmac. exp. Ther.* **199,** 553–564 (1976).
9. Kupfermann, I. *A. Rev. Physiol.* **42,** 629–641 (1980).
10. Kato, E. & Kuba, K. *J. Physiol., Lond.* **298,** 271–283 (1980).
11. Kuba, K. & Koketsu, K. *Prog. Neurobiol.* **11,** 77–169 (1978).
12. Del Castillo, J. & Katz, B. *J. Physiol., Lond.* **124,** 560–573 (1954).
13. Christ, D. D. & Nishi, S. *J. Physiol., Lond.* **213,** 107–117 (1971).
14. Kebabian, J. W. & Greengard, P. *Science* **174,** 1346–1349 (1971).
15. Cramer, H., Johnson, D. G., Hanbauer, I., Silberstein, S. D. & Kopin, I. J. *Brain Res.* **53,** 97–105 (1973).
16. Brown, D. A., Caulfield, M. P. & Kirby, P. J. *J. Physiol., Lond.* **290,** 441–451 (1979).
17. Kobayashi, K., Kuroda, Y. & Yoshioka, M. *J. Neurochem.* **36,** 86–91 (1981).
18. Shimahara, T. & Tauc, L. *J. Physiol., Paris* **74,** 515–519 (1978).
19. Klein, M. & Kandel, E. R. *Proc. natn. Acad. Sci. U.S.A.* **75,** 3512–3516 (1978).
20. Tada, M., Yamamoto, T. & Tonomura, Y. *Physiol. Rev.* **58,** 1–79 (1978).
21. Libet, B., Kobayashi, H. & Tanaka, T. *Nature* **258,** 155–157 (1975).
22. Kobayashi, H., Hashiguchi, T. & Ushiyama, N. S. *Nature* **271,** 268–270 (1978).

Functional EGF receptors are present on mouse embryo tissues

Eileen D. Adamson*, M. Jane Deller† & Joseph B. Warshaw‡

*La Jolla Cancer Research Foundation, 2945 Science Park Road,
La Jolla, California 92037, USA
†Department of Zoology, South Parks Road, Oxford University,
Oxford OXI 3PS, UK
‡Division of Perinatal Medicine, Department of Pediatrics and
Obstetrics and Gynecology, Yale University School of Medicine,
New Haven, Connecticut 06510, USA

Epidermal growth factor (EGF) is a small polypeptide (6,045 molecular weight) well recognized as a mitogen affecting a variety of cell types (see refs 1–3 for reviews). Its mitogenic effect is mediated through specific high-affinity receptors on the plasma membrane, and after ligand binding, some (but not all) cells respond in a variety of ways which lead to either increased growth rate or differentiation. Although the initial description of EGF by Cohen[4] demonstrated a maturational influence on eye-opening and tooth eruption in newborn mice, there has been little investigation of the effect of EGF on developing fetal tissues. Nexo *et al.*[5] have provided evidence that EGF and its receptors are present in mouse embryos as early as day 13 of gestation, especially in the secondary palate where EGF may have a role in palate closure, and EGF receptors have also been described in the human placenta[6] and in fetal rabbit lung[7]. Thus, the limited data available suggest that EGF may play a part in fetal growth and development of specialized function. Here, we describe the binding and mitogenic activity of EGF both *in vitro* and *in vivo* in several fetal mouse tissues.

Tissues dissected from embryos at days 11–18 of gestation (day 1 being the day that a copulation plug is found) were subjected to ^{125}I-EGF *in vitro* binding studies (Table 1). For most tissues the time of maximum concentration of EGF binding sites seems to be around the 14th day of gestation; however, this is only a rough estimate because different stages of maturation are encountered within one uterus. Amnion tissue showed the greatest binding activity (expressed per mg protein), followed by, in decreasing order, lungs, trophoblast, limbs, visceral yolk sac (VYS), liver, brain, parietal endoderm and heart.

In the standard assay described by us previously[8], tissues were exposed for 90 min at room temperature to tracer amounts of radioiodinated EGF in the presence or absence of a 500-fold excess of unlabelled EGF. Although nonspecific binding of whole tissue was higher than that found for cultured cells[8] (see Fig. 1*c* and Table 1 legend), specific binding was still readily measurable. Three processes seem to contribute to these values: (1) specific binding (that which is displaced by a large excess of unlabelled EGF); (2) endocytosis of receptor–ligand complexes (the inclusion of 0.1% sodium azide in the binding medium reduced the bound counts by 21% in amnion and by 25% in VYS and brain); (3) nonspecific binding and endocytosis detected by ^{125}I-EGF counts which are not displaced in the presence of cold EGF.

When cells were isolated from 15-day embryonal tissues using trypsin and then grown in monolayer cultures, the nonspecific binding was < 2% of the specifically bound counts. This suggests that for whole tissues some ^{125}I-EGF is trapped, whereas trypsin-treated cells have better access to the binding medium as well as a less 'sticky' surface.

An additional difficulty in using whole tissues is that of the accessibility of the tracer to the deeper-lying cells. We therefore made crude cell membrane preparations of each tissue and measured binding to aliquots of these preparations (Table 1). Although we used fewer samples than for whole tissue, the results with membrane preparations confirmed the whole tissue results.

In the experiments described above, parietal endoderm (cells plus basement membrane) showed low but significant binding of ^{125}I-EGF. However, this tissue is difficult to dissect entirely free of trophoblast cells, which bind higher levels (see Table 1). In addition, autoradiographic studies showed that Reichert's membrane can bind EGF. Therefore, we prepared single-cell suspensions of parietal endoderm cells by incubating yolk sacs with trypsin (0.25% w/v) and EDTA (0.1 mM) in phosphate-buffered saline (PBS) and allowed these cells to attach to plastic dishes (10^5 cells per 35-mm dish). Cells were incubated in

Table 1 *In vitro* binding of ^{125}I-EGF to mouse embryonic tissues

Tissue	Day of gestation 11th and 12th	13th and 14th	15th and 16th	17th and 18th
Placenta membranes*	0.46 (3)	1.33 (37)	2.70 ± 2.1 (8)	0.35 (4)
Trophoblast	0.52 ± 0.05 (5)	2.34 ± 0.07 (7)	—	—
PYS endoderm*	0.25 ± 0.02 (4)	0.32 ± 0.02 (10)	—	—
VYS	1.82 ± 0.02 (5)	0.73 ± 0.03 (11)	1.02 ± 0.04 (12)	0.88 ± 0.03 (5)
Amnion	14.6 ± 5.9 (5)	21.5 ± 5.3 (11)	15.7 ± 3.5 (12)	9.1 ± 4.0 (5)
Amnion membranes*	ND	78.7 (2)	67.3 (1)	13.4 (1)
Limbs	—	1.7 ± 1.2 (4)	2.2 (1)	1.26 (1)
Lungs	—	4.26 (2)	0.44 ± 0.16 (4)	0.86 (1)
Lung membranes*	—	2.2 (1)	1.08 ± 0.41 (8)	0.66 ± 0.19 (5)
Liver	—	0.66 ± 0.18 (5)	0.08 (4)	0.23 (3)
Liver membranes*	—	0.96 ± 0.82 (4)	0.44 ± 0.24 (8)	0.52 (3)
Brain	—	0.48 ± 0.13 (4)	1.06 (3)	0.7 (1)
Brain membranes*	—	6.8 (3)	0.36 ± 0.10 (8)	0.43 ± 0.02 (5)
Heart	—	0.2 (3)	0.9 (2)	ND
Heart membranes*	—	ND	1.63 ± 0.97 (7)	2.23 (3)

EGF (Collaborative Research) was iodinated and purified by passage through Sephadex G-50 as described by Rees *et al.*[8]. Embryos were dissected in PBS (solution A)[22] and washed in the same solution. Individual or pooled tissues were incubated in 24-well microtitre plates or in plastic tubes with 1 ml 'binding' buffer. Binding buffer contained 1 mg ml^{-1} bovine serum albumin in Earle's balanced salt solution and 15 mM HEPES adjusted to *p*H 7.4 with 1 M NaOH. Iodinated EGF was added to 8×10^{-11} M (0.4 ng, 70,000 c.p.m.). After gentle agitation for 90 min at room temperature, tissues were washed three times in the same buffer at 0 °C (minus EGF) and twice in ice-cold PBS. Control samples were incubated similarly except for the addition of unlabelled EGF at 500 times the concentration of the labelled form. Each washed tissue was solubilized in 0.5–1.0 ml 0.2 M NaOH at 37 °C for 30 min and counted for 1 min in a Wallac γ-counter. Samples were assayed for protein by the Coomassie blue method[23]. Similar assays for ^{125}I-EGF binding were carried out on crude membrane preparations. Various tissues were homogenized in glass/glass 1.0 ml homogenizers with 0.5 ml 0.01 M Tris-HCl, 0.01 M NaCl and 3 mM $MgCl_2$, *p*H 8.0. The homogenates were centrifuged at 100,000*g* for 30 min and the pellets resuspended in the binding buffer described above. Aliquots were tested for specific binding of ^{125}I-EGF and for protein content as above. The membranes were washed by suspension and centrifugation at 10,000*g* for 10 min or 5,000*g* for 20 min. Nonspecific binding of ^{125}I-EGF to embryonic tissues varied from 20% of total counts bound in amnion to 51.7% in placenta; in the crude membrane preparations, nonspecific binding varied from 26% for brain to 76% in limbs. For each tissue in each experiment the results were expressed as c.p.m. ^{125}I-EGF bound per mg protein minus c.p.m. ^{125}I-EGF bound in the presence of 500-fold excess of EGF per mg tissue protein divided by the total c.p.m. added to the incubation mixture and multiplied by 100. At least six such samples of each tissue were examined in each experiment (three reaction samples and three controls). ND, not done; PYS, parietal yolk sac; VYS, visceral yolk sac. The number of experiments is shown in parentheses.

*Crude plasma membrane preparations. parietal yolk sac endoderm cells plus Reichert's membranes are also contaminated with trophoblast cells.

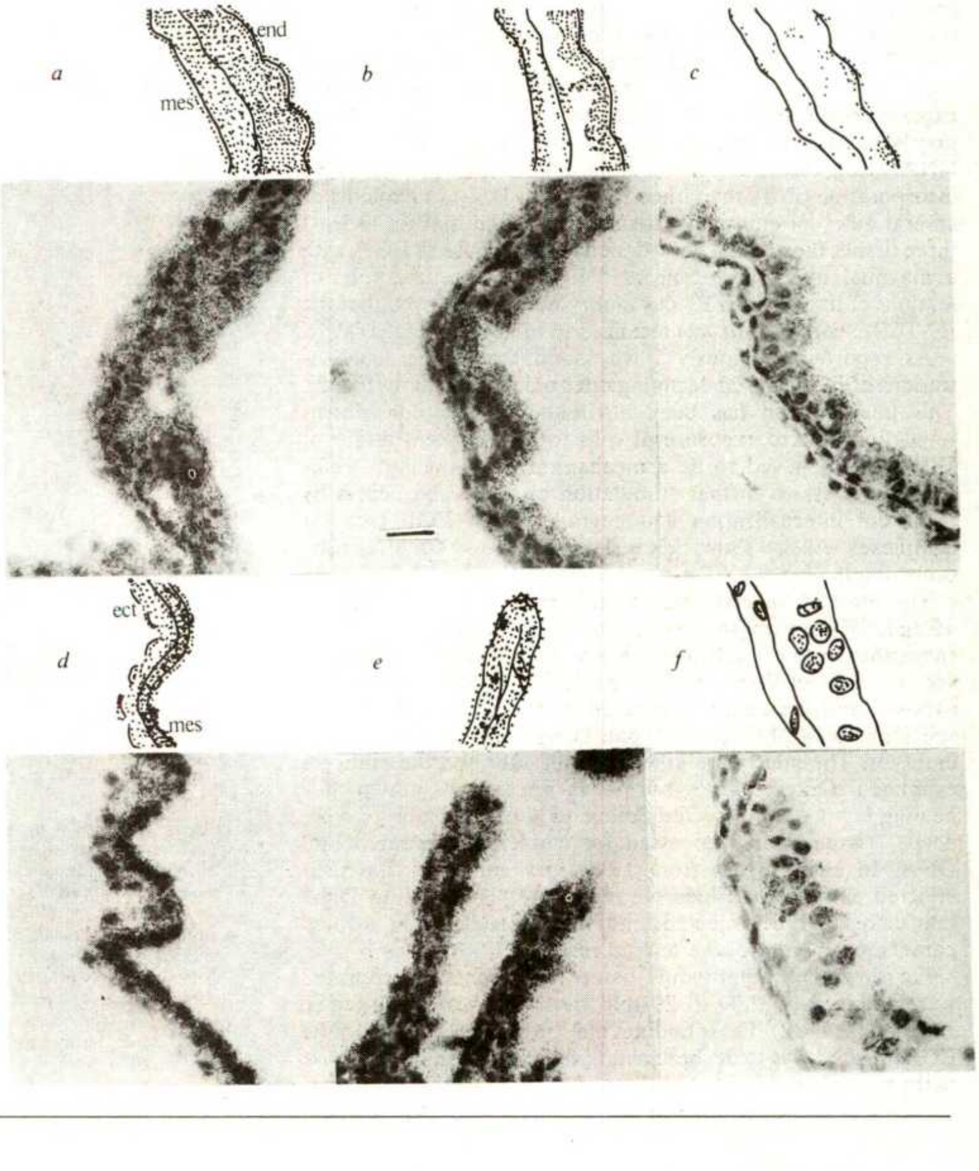


Fig. 1 Autoradiographs of 14-day mouse embryo tissues to show ^{125}I-EGF binding and ^{3}H-TdR incorporation into the nucleus. For binding studies, tissues were incubated for 90 min: *a*, *d*, in 0.3 ml binding medium containing 250,000 c.p.m. ^{125}I-EGF; *b*, *e*, the same plus 0.1% sodium azide; *c*, in 0.3 ml binding medium containing 250,000 c.p.m. ^{125}I-EGF plus 500 ng unlabelled EGF. The tissues were washed six times in binding buffer and were fixed in phosphate-buffered 1% (w/v) formaldehyde for 4–6 h at room temperature. Tissues were washed, dehydrated and embedded in paraffin wax, sectioned at 8 μm and slices were dried on to slides. After dewaxing, slides were exposed to Kodak AR-10 stripping film for 6 weeks before being developed. *a–c*, VYS; *d*, *e*, amnion. Photographic reproduction of these sections is difficult and therefore diagrams have been drawn above the autoradiographs to indicate the relative approximate density of grains. Note that the VYS endoderm still takes up labelled EGF in the presence of a large excess of cold EGF (*c*). This nonspecific labelling probably occurs partly by endocytosis, a process which is characteristic of VYS endoderm[24]. *d*, Shows that one cell layer of the amnion (tentatively identified as mesoderm) can bind greater amounts of EGF than the other cell layer. *f*, Autoradiograph of VYS incubated *in vitro* with EGF (5 ng ml^{-1}) for 16 h and then with ^{3}H-TdR (1 μCi ml^{-1}) for 4 h before being fixed and processed as above. End, endoderm; ect, ectoderm; mes, mesoderm. Scale bar, 20 μm.

Table 2 Incorporation of ^{3}H-TdR into extra-embryonic tissues *in vivo*

	c.p.m. per mg protein ($\times 10^{-3}$) Placenta	VYS	Amnion
EGF injected	1.6 (7)	15.0 (7)	133.0 (4)
PBS injected	5.6 (8)	443.0 (7)	304.8 (5)
Ratio PBS/EGF	3.5	29	2.3

The values represent averages of three different experiments on four different females, with the number of embryos examined in parentheses. Experimental method is described in the text and the tissues were processed as described in Fig. 2 legend.

medium which had been 'conditioned' by PYS-2 teratocarcinoma cells[9] and filtered through 0.2-μm Millipore filters. Parietal endoderm cells attached well in these conditions and were incubated for 24 h to allow them to recover from any loss of trypsin-sensitive EGF-binding activity. Similar cell preparations were made from amnions to act as a standard cell type with high binding activity. Binding assays[8] on cells attached to dishes were performed essentially as described for suspended tissues. Little or no ^{125}I-EGF binding was detected in parietal endoderm, whereas 4.1% specific binding per 10^6 cells was detected in amniotic cells.

Visceral yolk sac consists of two layers, endoderm cells underlaid by basement membrane, and mesoderm consisting of several different kinds of cells including endothelial, fibroblastic and haematopoietic cells. Amnion has an ectodermal and a mesodermal layer. We examined the distribution of EGF-binding activity in the layers of VYS and amnion by autoradiographs prepared as described in Fig. 1 legend, and found that the greatest concentration of specific labelling occurred in the endoderm of the VYS (Fig. 1*a–c*) and in one of the two layers of the amnion (Fig. 1*d*), although all cell types had binding activity.

EGF has a wide range of effects on different cell types *in vitro*[5,10–15]. The response of a target cell is usually determined experimentally through the mitogenic activity of EGF which may be measured by increased uptake of tritiated thymidine into DNA. Using assays described in Fig. 2 legend, we measured the incorporation of ^{3}H-thymidine (TdR) into DNA of explants of several different embryonic tissues, and found that for at least three tissues there was a dose-dependent response to EGF, with a maximal effect at ~ 5 ng ml^{-1} EGF. Autoradiographs of sections of tissues from 13-day embryos clearly showed that the ^{3}H-TdR incorporation was localized in the nuclei (Fig. 1*f*).

As reported previously[16], we found that higher concentrations of EGF caused declining rates of incorporation (Fig. 2). This phenomenon has been attributed to receptor 'down-regulation' due to exposure of cells to high concentrations of EGF. It is believed to be a mechanism for reducing a cell's responsiveness to further stimulation by EGF and occurs by persistent internalization and degradation of EGF–receptor complexes. As we show below, such a process could also be occurring *in vivo*.

The uteri of anaesthetized female mice were exposed and 10 ng EGF per embryo were injected into the amniotic cavities. In another female, control embryos were injected with the same volume (2μl) of PBS, and 16–18 h later, the uteri were again exposed and the embryos injected with 2 μl ^{3}H-TdR (1 μCi, specific activity 24 Ci mmol^{-1}) plus 10 ng EGF for experimental embryos. The mice were killed 90 min later and the embryos examined. Only embryos normal in appearance and with a beating heart were dissected (about 15% of the embryos were dead). Tissues were processed for counts incorporated into DNA. In eight tissues from 13–14-day embryos, the PBS-injected embryos had incorporated more ^{3}H-TdR into DNA than had the EGF-injected embryos, a result which seemed paradoxical in view of the *in vitro* results.

For three extra-embryonic tissues the counts incorporated were reduced by 2.3- to 29-fold those in the EGF-injected embryos (Table 2). These findings may be simply explained if the EGF injected was toxic, but general toxicity is unlikely because of the variable ratio of reduced incorporation in different tissues and because even when massive doses of EGF were injected, similar results were obtained. Also arguing against toxicity are the parallel results achieved *in vivo* and *in vitro* at similar concentrations of EGF and ^{3}H-TdR. Instead, we hypothesize that EGF receptors may be down-regulated by prolonged exposure of the embryo to exogenous EGF.

EGF is present in mouse serum, urine and milk[17] and also occurs in human amniotic fluid (at 16–18 weeks gestation, 0.9–8.8 ng ml^{-1}, ref. 18). EGF-like material is extractable from mouse fetuses at levels ranging from 0.6 to 15 ng per 14-day embryo[5] although its level in amniotic fluid has not been reported. In our experiments the addition of ≥ 10 ng of EGF per embryo seems to be sufficient to down-regulate the receptors. Such a conclusion is supported by the data in Fig. 2.

EGF has several roles other than mitogenic effects and we therefore cannot expect a simple relationship between EGF binding and thymidine incorporation, although for fibroblasts and many other cell types *in vitro*, this relationship largely holds. For embryonic tissues, it can also be speculated that EGF stimulates the differentiation of these tissues as it does in lung development in sheep[19], rabbits[7] and chicks[20].

In conclusion, most embryonic tissues bind EGF. Binding correlates with ^{3}H-TdR incorporation *in vitro* and evidence is provided that down-regulation of EGF receptors occurs *in vitro* and perhaps *in vivo*. These are some of the essential criteria which identify specific receptors for EGF[21]. The presence of EGF binding activity may serve to distinguish between extra-embryonic endoderm cell types, as visceral endoderm but not parietal endoderm binds EGF. This property will be useful for characterizing the endoderm cell products of teratocarcinoma EC cells.

We thank Professor R. L. Gardner for inviting J.B.W. to work in his laboratory, and Drs C. F. Graham and D. A. Mercola for useful discussions and comments on the script. This work was supported by grants to Dr C. F. Graham from the MRC, the Cancer Research Campaign and NIH grant 08293.

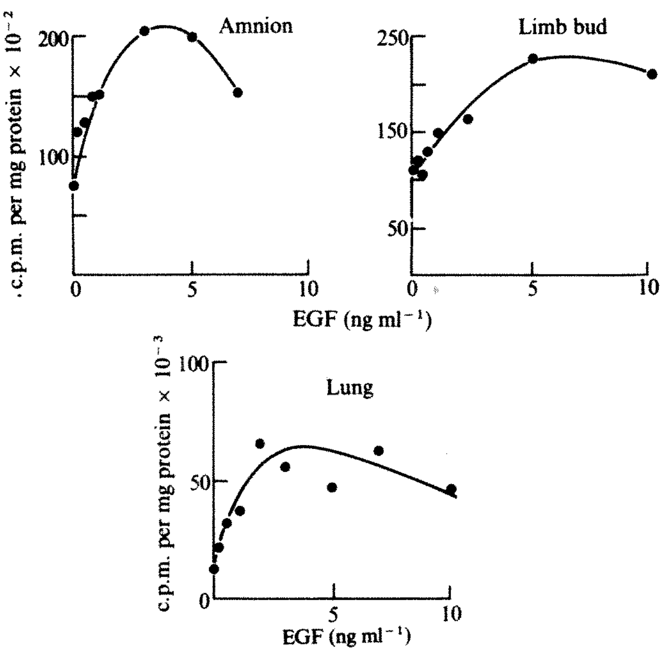


Fig. 2 The stimulation of DNA synthesis by EGF in mouse embryo tissues incubated *in vitro*. Tissues from 15-day embryos were dissected and incubated in culture media containing 10% dialysed fetal calf serum and concentrations of EGF varying from 0 to 10 ng ml^{-1} and 1 μCi ml^{-1} ^{3}H-TdR (24 Ci mmol^{-1}) for 16 h at 37 °C in a gassed incubator. The tissues were washed three times in PBS and then unincorporated label was removed by solubilization of tissues in 0.5 ml 0.5 M NaOH and subsequent precipitation by 0.5 ml 20% (w/v) trichloroacetic acid. After washing twice by resuspension and centrifugation the pellet was assayed for protein content and an aliquot tested for incorporated tritium radioactivity. The results are expressed as c.p.m. incorporated per mg protein in the sample for each EGF concentration. The graph represents one of three similar experiments for three different tissues but other tissues gave similar results.

Received 24 November 1980; accepted 14 April 1981.

1. Carpenter, G. & Cohens, S. *A. Rev. Biochem.* **48,** 193–216 (1979).
2. Hollenberg, M. D. *Vitam. Horm.* **37,** 69–110 (1979).
3. Adamson, E. D. & Rees, A. R. *Molec. cell. Biochem.* **34,** 129–152 (1981).
4. Cohen, S. *Devl Biol.* **12,** 394–407 (1965).
5. Nexo, E., Hollenberg, M. D., Figueroa, A. & Pratt, R. M. *Proc. natn. Acad. Sci. U.S.A.* **77,** 2782–2785 (1980).
6. O'Keefe, E., Hollenberg, M. D. & Cuatrecasas, P. *Archs Biochem. Biophys.* **164,** 518–526 (1974).
7. Catterton, Z. *et al. Pediat. Res.* **13,** 104–108 (1979).
8. Rees, A. R., Adamson, E. D. & Graham, C. F. *Nature* **281,** 309–311 (1979).
9. Lehman, J. M., Speer, W. C., Swartzendruber, D. E. & Pierce, G. B. *J. cell. Physiol.* **84,** 13–28 (1974).
10. Johnson, L. K., Baxter, J. D., Vlodavsky, I. & Gospodarowicz, D. *Proc. natn. Acad. Sci. U.S.A.* **77,** 394–398 (1980).
11. Tonelli, Q. J. & Sorof, S. *Nature* **285,** 250–252 (1980).
12. Covelli, I., Rossi, R., Mozzi, R. & Frati, L. *Eur. J. Biochem.* **27,** 225–230 (1972).
13. Carpenter, G., King, L. & Cohen, S. *Nature* **276,** 409–410 (1978).
14. Barnes, D. & Colowick, S. P. *J. cell. Physiol.* **89,** 633–640 (1976).
15. Hollenberg, M. D. & Cuatrecasas, P. *J. biol. Chem.* **250,** 3845–3853 (1975).
16. Fox, F. & Das, M. *J. supramolec. Struct.* **10,** 199–214 (1979).
17. Bynny, R. L., Orth, D. N., Cohen, S. & Doyne, R. S. *Endocrinology* **95,** 776–782 (1974).
18. Ladda, R. L., Bullock, L. P., Gianopoulos, T. & McCormick, L. *Analyt. Biochem.* **93,** 286–294 (1979).
19. Sundell, H. *et al. Pediat. Res.* **9,** 371A (1975).
20. Goldin, G. V. & Opperman, L. A. *Eur. J. cell. Biol.* **22,** 388A (1980).
21. Hollenberg, M. D. & Cuatrecasas, P. in *The Receptors* Vol. I (ed. O'Brien, R. D.) 193–213 (Plenum, New York, 1979).
22. Dulbecco, R. & Vogt, M. *J. exp. Med.* **99,** 167–182 (1954).
23. Sedmak, J. J. & Grossberg, S. E. *Analyt. Biochem.* **79,** 544–552 (1977).
24. Wild, A. E. *Phil. Trans. R. Soc.* **B271,** 395–410 (1975).

(−)Maackiain acetate specifically inhibits different forms of aryl hydrocarbon hydroxylase in rat and man

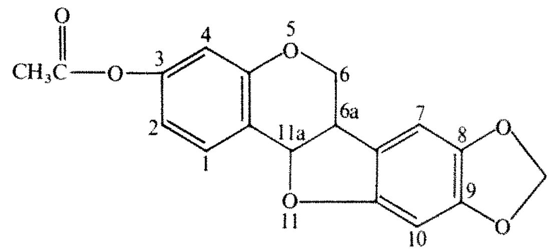
H. V. Gelboin*, Donna West*, E. Gozukara*, S. Natori†, Minako Nagao‡, & T. Sugimura‡

* Laboratory of Molecular Carcinogenesis, National Cancer Institute, Bethesda, Maryland 20205, USA
† National Institute of Hygiene Science, Kamiyoga, Setagayaku, Tokyo 158, Japan
‡ Division of Biochemistry, National Cancer Research Institute, Tsukiji, Chuo-ku, Tokyo, 104, Japan

The mixed-function oxidases[1], which contain different forms of cytochrome P450, are the critical enzymes which metabolize a wide variety of drugs, carcinogens and environmental chemicals, as well as steroids and other endogenous metabolites[2]. The metabolism of these chemicals may result either in their detoxification or in their conversion to toxic, mutagenic and carcinogenic metabolites. Modulation of the mixed-function oxidase, aryl hydrocarbon hydroxylase (AHH), by either inhibitors or inducers strongly affects carcinogenesis induced by polycyclic hydrocarbons[3,4]. We reported earlier that 7,8-benzoflavone is a potent inhibitor of the form of AHH that is induced by polycyclic hydrocarbons but either weakly inhibits or stimulates the form of AHH derived from the livers of phenobarbital-treated or untreated rats[5,6]. We now report that the pterocarpan (−)maackiain acetate exhibits unique specificity that is the reverse of 7,8-benzoflavone, that is, it has strong inhibitory activity towards the form of AHH found in human liver or in untreated or phenobarbital-treated rats and weak inhibitory or stimulatory activity towards the AHH induced by polycyclic hydrocarbons in rat tissues or by cigarette smoking in human placenta. Thus, we conclude that (−)maackiain acetate and 7,8-benzoflavone inhibit different forms of AHH.

The structure of (−)maackiain acetate is shown below. (−)Maackiain is found in leguminous plants[7,8] and was isolated from the heartwood constituents of *Sophora japonica* L. and converted to the acetate as described by Suginome[7,9]. The physical characteristics of the material were as described previously[7].

Three to five male Sprague–Dawley rats (120–160 g) were injected intraperitoneally with either 1.3 mg methylcholanthrene per 0.3 ml corn oil per day for 3 successive days or 80 mg phenobarbital per 80 kg body weight per day for 3 days. The rats were killed and microsomal preparations prepared from liver, kidney and lung by published procedures[6]. Human liver was obtained within 6 h after the death of an accident victim and microsomes were prepared by a similar method to that described for rat liver. Placental tissue was obtained immediately post partum from women who had smoked at least 20 cigarettes per day for an extended period. The placentas were chilled on ice and microsome pellets obtained by essentially the same procedure as for rat liver. Protein content was determined by the method of Lowry *et al.*[10]. AHH activities of the various preparations were determined during a 10-min incubation with liver microsomes and for a 30-min incubation with microsomes from lung, kidney and placenta[11]. The concentration of the substrate benzo(*a*)pyrene was 10^{-4} M in all experiments. Metabolite analysis of benzo(*a*)pyrene metabolism was analysed by HPLC using methods we have described previously[12].



(−) Maackiain acetate

Table 1 contrasts the effects of (−)maackiain acetate and 7,8-benzoflavone on AHH activity. At 10^{-4} M (−)maackiain acetate, a concentration equimolar to the substrate concentration, AHH activity in liver microsomes was inhibited by 95% in control rats and 76% in phenobarbital-treated rats, suggesting that the phenobarbital inducible enzymes may be rather less susceptible to (−)maackiain acetate inhibition than the enzyme forms in control rats. (−)Maackiain acetate either stimulated or had no inhibitory effect on the AHH of liver microsomes from methylcholanthrene-induced rats. In contrast, 7,8-benzoflavone strongly inhibits the AHH from methyl-

Table 1 Differential effects of (−)maackiain acetate and 7,8-benzoflavone on AHH of liver, lung, and kidney from normal, phenobarbital- and methylcholanthrene-treated rats

Expt	Tissue	*In vivo* treatment	Inhibitor (10^{-4} M)	AHH (pmol per min per mg protein)	% Control
1	Liver	Control	None	581	—
			MA	31	5
			7,8-BF	1,030	178
		Phenobarbital	None	862	—
			MA	210	24
			7,8-BF	1,300	151
		Methyl-cholanthrene	None	1,884	—
			MA	2,982	158
			7,8-BF	636	34
2	Lung	Control	None	8	—
			MA	3	37
			7,8-BF	7	87
		Methyl-cholanthrene	None	107	—
			MA	102	95
			7,8-BF	40	37
	Kidney	Control	None	23	—
			MA	1	4
			7,8-BF	17	74
		Methyl-cholanthrene	None	101	—
			MA	56	55
			7,8-BF	39	39

Experiments 1 and 2 were performed on different sets of rats. MA, (−)maackiain acetate; 7,8-BF, 7,8-benzoflavone.

0028–0836/81/260659—03$01.00

Table 2 Differential effects of (−)maackiain acetate and 7,8-benzoflavone on AHH of normal human liver and placenta from cigarette-smoking women

Tissue	Inhibitor (M)	AHH (pmol per min per mg protein)	% Control
Liver	None	11.0	—
	MA (10^{-4})	3.5	3.2
	MA (10^{-5})	5.5	50
	7,8-BF (10^{-5})	39.0	354
Placenta	None	15.6	—
	MA (10^{-4})	20.4	131
	MA (10^{-5})	20.6	132
	7,8-BF (10^{-4})	3.5	22
	7,8-BF (10^{-5})	8.1	39

cholanthrene-induced rats but has little or no inhibitory effect or stimulates the AHH from controlled or phenobarbital treated rats[5,6,13,14].

(−)Maackiain acetate and 7,8-benzoflavone also distinguish between two forms of the enzyme in lung and kidney. With lung microsomes, (−)maackiain acetate strongly inhibits AHH from control rats and does not inhibit the methylcholanthrene-induced enzyme, whereas 7,8-benzoflavone inhibits only the latter. The 7,8-benzoflavone inhibition of lung AHH from methylcholanthrene-induced rats has been reported to be only slightly greater than that from control rats[13]. With kidney microsomes, the (−)maackiain acetate inhibits both the control and methylcholanthrene-induced AHH but the latter to a much lesser extent. These data indicate that the composition of AHH present in the lung and kidney of untreated rats is different from that in the methylcholanthrene-inducible tissues.

Table 3 Characteristics of inhibition of (−)maackiain acetate and 7,8-benzoflavone in rat liver microsomes

		Microsome source		
	Inhibitor	Control	Phenobarbital	Methylcholanthrene
ID50	(−)MA	1×10^{-6} M	2×10^{-5} M	—
	7,8-BF	—	—	2×10^{-5}
K_m	None	7×10^{-6}	—	1.2×10^{-5}
	MA (10^{-5} M)	8×10^{-6}	—	—
	7,8-BF (10^{-5} M)	—	—	1.2×10^{-5}
V_{max}	None	1,720	—	1,000
	(−)MA (10^{-5} M)	200	—	—
	7,8-BF (10^{-4} M)	—	—	500
K_I	MA	(1.6×10^{-6})	—	—
	7,8-BF	—	—	8.8×10^{-6}

The ID_{50} is the molar concentration that gives 50% inhibition and was calculated from direct measurement of inhibition at different concentrations of inhibitor. The K_I was calculated from Michaelis–Menten curves. V_{max}, units pmol per min per mg protein.

Table 2 shows the effects of (−)maackiain acetate and 7,8-benzoflavone on normal human liver, presumably uninduced for AHH, and the AHH of human placenta induced by heavy cigarette smoking, thus representing, respectively, control and polycyclic hydrocarbon-induced human tissues. (−)Maackiain acetate strongly inhibits human liver AHH and slightly stimulates the AHH of the induced placenta. AHH is essentially undetectable in the placentas of non-smoking women. In contrast, 7,8-benzoflavone caused a large stimulation of AHH in human liver and strongly inhibited inducible AHH of human placenta. A similar stimulation of control and phenobarbital-induced rat liver AHH[13,14] and human liver AHH[15] has been reported previously.

Table 3 shows some of the characteristics of (−)maackiain acetate and 7,8-benzoflavone with respect to inhibitory potency and effect on the kinetic characteristics of the AHH. The ID_{50} of (−)maackiain acetate for control rat liver AHH is 10^{-6} M; that of 7,8-benzoflavone for methylcholanthrene-inducible rat liver AHH is 10^{-5} M. Each inhibitor reduces the V_{max} and has relatively little effect on the K_m of the AHH. Both inhibitors are non-competitive with substrate and (−)maackiain acetate is about 10-fold more potent for the control AHH than 7,8-benzoflavone for the polycyclic hydrocarbon-induced AHH.

Table 4 shows an HPLC analysis of the effect of (−)maackiain acetate on the formation of individual benzo(*a*)pyrene metabolites. Our earlier studies showed that 7,8-benzoflavone reduced the amount of each metabolite formed by methylcholanthrene-induced microsomes to about the same extent[16]. (−)Maackiain acetate drastically inhibits the formation of the 7,8-diol, 9,10-diol, 9-OH, 3-OH and 7-OH, but has a negligible effect on the formation of the 4,5 diol, suggesting that this metabolite may be formed by a different mixed-function oxidase from that metabolizing the formation of the other metabolites. The amount of quinones formed is reduced to a lesser extent than the major metabolites, which suggests either an alternative mixed-function oxidase for their formation or possibly their formation by a mechanism other than via the mixed-function oxidases.

The multiple forms of the cytochrome P450-containing mixed-function oxidases display different catalytic activities in the metabolism of various drugs[1,17] and exhibit both substrate and product regio- and stereospecificity for the polycyclic hydrocarbon benzo(*a*)pyrene, a prototype carcinogen that is metabolized by this enzyme system[12]. Metabolic pathway choice for a specific xenobiotic may thus be largely determined by the type and amount of the mixed-function oxidase engaging its metabolism.

Of a series of derivatives of 7,8-benzoflavone and other flavonoids, several were found either to inhibit the methylcholanthrene-induced AHH form only or to inhibit both the control and methylcholanthrene-induced forms[13,18]. None of the many flavonoids tested exhibited the strong inhibitory activity specific for the control enzyme that is shown by (−)maackiain acetate. The unusual inhibitory specificity of (−)maackiain acetate may make it an extremely valuable tool for probing the enzymology and active catalytic sites of the different

Table 4 HPLC analyses of (−)maackiain acetate effect on benzo(*a*)pyrene metabolism in liver microsomes from control and MC-treated rats

	Control microsomes			MC-microsomes		
Metabolite formed	No inhibitor (pmol per min per mg protein)	(−)Maackiain acetate	% Control	No inhibitor (pmol per min per mg protein)	(−)Maackiain acetate	% Control
9,10-diol	33.2	6.0	18	572	608	106
4,5-diol	25.0	24.0	96	217	289	133
7,8-diol	8.6	0.7	8	326	368	113
9-OH	5.8	—	<1	241	386	160
7-OH	—	—	—	18	—	—
3-OH	222.1	17.8	8	1,261	1,471	116
Quinones	39.3	98.0	249	460	455	101
Total	334.0	146.5	43	3095	3577	729

forms of cytochrome P450, and the factors that determine metabolic pathway choice, and may thus help us to understand the relationship between drug and carcinogen activation and detoxification. It is also possible that individual or combinations of potent inhibitors of the mixed-function oxidases, such as (−)maackiain acetate and 7,8-benzoflavone, may be powerful natural inhibitors or modulators of chemical carcinogenesis.

Received 15 December 1980; accepted 6 April 1981.

1. Coon. M. J. & Vatsis, K. P. in *Polycyclic Hydrocarbons and Cancer* Vol. 2. (eds Gelboin, H. V. & Ts'o, P. O. P.) 335–360 (Academic, New York, 1978).
2. Conney, A. H. *Pharmac. Rev.* **19,** 317–366 (1967).
3. Gelboin, H. V. *Physiol. Rev.* **60,** 1107–1166 (1980).
4. Wattenberg, L. W. *Cancer Res.* **26,** 1520–1526 (1966).
5. Diamond, L. & Gelboin, H. V. *Science* **166,** 1023–1025 (1969).
6. Wiebel, F. J. Leutz, J. C., Diamond, L. & Gelboin, H. V. *Archs Biochem. Biophys.* **144,** 78–86 (1971).
7. Suginome, H. *Bull. chem. Soc. Japan* **39,** 1529–1534 (1966).
8. Shibata, S. & Nishikawa, Y. *Chem. pharm. Bull., Tokyo,* **11,** 167–177 (1963).
9. Suginome, H. *J. org. Chem.* **24,** 1655–1662 (1959).
10. Lowry, O. H., Rosebrough, N. A. Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193,** 265–275 (1951).
11. Nebert, D. W. & Gelboin, H. V. *J. biol. Chem.* **243,** 6242–6261 (1968).
12. Deutsch, J. *et al. Proc. natn. Acad. Sci. U.S.A.* **75,** 3123–3127 (1978).
13. Wiebel, F. J., Gelboin, H. V., Buu-Hoi, N. P., Stout, M. G. & Burnham, W. S. in *Chemical Carcinogenesis* Pt A (eds Ts'o, P. & Di Paolo, J. A.) 249–270 (Dekker, New York, 1974).
14. Wiebel, F. J. & Gelboin, H. V. *Biochem. Pharmac.* **24,** 1511–1515 (1975).
15. Kapitulnik, J., Poppers, P. J., Buening, M. K., Fortner, J. G. & Conney, A. H. *Clin. Pharmac. Ther.* **22,** 475–485 (1977).
16. Selkirk, J. K., Croy, R. G., Roller, P. P. & Gelboin, H. V. *Cancer Res.* **34,** 3474–3480 (1974).
17. Ryan, D. E., Thomas, P. E., Korzeniowski, D. & Levin, W. *J. biol. Chem.* **254,** 1365–1374 (1979).
18. Nesow, S. *J. med. Chem.* **22,** 1244–1247 (1979).

Idiosyncratic platelet responses to human tumour cells

Eva Bastida, Antonio Ordinas & G. A. Jamieson

American Red Cross, Blood Services Laboratories, Bethesda, Maryland 20014, USA

Platelets are thought to be involved in two important steps in the dissemination of cancer: (1) in their interaction with cells released from the primary tumour to form a mixed platelet–tumour cell embolus and (2) in the subsequent attachment or entrapment of this embolus at the vessel wall and the resulting extravasation of the tumour cell and growth of the secondary tumour[1]. We have recently used the Baumgartner perfusion chamber[2] to evaluate the attachment phase of platelet–tumour cell interactions at subendothelial surfaces[3]. We have found that tumour cells which caused platelet aggregation (Hut 20 line, large cell carcinoma of the lung) were incorporated into mixed platelet–tumour cell thrombi on basement membrane or collagen microfibrils, while those tumour cells that did not cause platelet aggregation (A549 line, epithelial lung carcinoma), were not incorporated in this way. We have now found differences between different donors in the ability of their platelets to aggregate in response to various human tumour cell lines. These different individual susceptibilities seem to be largely independent of the cell line involved and may have implications in the evaluation of mechanisms of human neoplastic disease.

The primary interaction of platelets and tumour cells has usually been studied by aggregometry[4–7] and a rough correlation has been established between the ability of tumour cells to aggregate platelets and their ability to induce lung metastases in mice[4]. The basis for this aggregation by tumour cells has not been determined but has generally been ascribed as being due to the release of ADP from platelets[4,5], to the thromboplastic activity of tumour cells[8] or to a lipoglycoprotein complex of the tumour cell surface[7]. This complex may be shed as microvesicles[4] and may be identical to the platelet aggregating material (PAM) which has been isolated from mouse tumour cells[6]. However, PAM cannot be isolated from tumour cells in all cases[7] and other mechanisms may be involved for other cell types.

Chelation of Ca^{2+} by collection of blood into citrate anticoagulant results in the irreversible loss of the ability of platelets to interact with tumour cells. In our experiments, platelet-rich plasma was prepared from blood collected in sodium heparin ($5\ U\ ml^{-1}$) from healthy volunteers who had taken no drugs for the previous 7 days: platelet counts (~300,000 per μl) and recoveries (~80%) were normal. Aliquots of washed tumour cells from each cell line were then added, to final concentrations given in Table 1, to an aliquot of platelet-rich plasma from each donor in the aggregometer (Chronolog, Broomall, Pennsylvania) and the extent of aggregation determined. Aggregation induced by 10 μM ADP was taken as 100% response. After culture of each cell line in standard conditions, cells were collected by decanting the culture medium, and the monolayers were washed twice with Hank's balanced salt solution (HBSS) followed by treatment for 5 min with HBSS free of Ca^{2+} and Mg^{2+} and containing 4 mM EGTA. The cell suspension was centrifuged at 800*g* for 10 min, the supernatant solution removed for testing and the cell pellets were washed twice with a solution of HBSS free of Ca^{2+} and Mg^{2+} but containing bovine serum albumin and $0.25\ mg\ ml^{-1}$ apyrase. Finally they were resuspended in the same solution but without apyrase. Cells were counted in a haemocytometer and viability determined by exclusion of Trypan blue. The range of viable cells was between 90 and 97%. None of the cell lines showed contamination with fibroblasts and the effects observed were not affected by treatment of the cultures with collagenase.

The results are summarized in Table 1. Individual donors were examined 2–4 times over a period of 4 months. Results were constant with regard to aggregation response in comparison with the 10 μM ADP control and lag times did not vary by more than ±10 s in repeat experiments. At a concentration of 10^5 cells ml^{-1} the U87MG (glioblastoma) line caused the aggregation of platelet-rich plasma (PRP) from each of the 16 donors examined. With Hut 20 (large cell lung carcinoma) cells at this concentration, six of the donors showed full platelet aggregation response with normal lag time (2–3 min), while the lag time to the onset of aggregation was prolonged in two other donors (5 and 9) and six failed to show aggregation. At 10^6 cells ml^{-1}, the Hut 28 (mesothelioma) line gave no aggregation with the same six donors who had shown zero aggregation response with Hut 20 at $10^5\ ml^{-1}$, and also with the two donors (5 and 9) who had shown a prolonged lag time to onset of aggregation. Donor 10, who had shown full aggregation with normal lag time with Hut 20 cells, did not give an aggregation response with HT 29 (adenocarcinoma). All donors showed a full aggregation response with Hut 20 at 10^6 cells ml^{-1}. For Hut 28 and SKNMC (neuroblastoma), the aggregation threshold was not reached until 5×10^6 cells ml^{-1} but, at this concentration, the same nine donors did not show an aggregation response, as was the case with the HT 29 line at 10^6 cells ml^{-1}. None of the donors responded to A549 (epithelial lung carcinoma) at tumour cell concentrations as high as $10^7\ ml^{-1}$.

The maximum aggregation response was not the same for each cell line: the U87MG, Hut 20 and HT 29 lines gave maximum responses, that is, equal to that observed with 10 μM ADP, whereas the responses of the SKNMC and Hut 28 lines were 60 and 50% respectively, of the ADP value. However, these values were reproducible and probably reflect differences in the sizes of the platelet–tumour cell aggregates in these two cases.

Donor 7 was unusual in showing reduced aggregation (~40%) with U87MG at 10^5 cells ml^{-1} with a prolonged lag time of 4.2 min and there was no response with any other cell line.

These results show, not surprisingly, that certain tumour cell lines do not cause platelet aggregation (A549) whereas others (U87MG) cause aggregation at all tumour cell concentrations examined. However, for several lines, there are critical tumour cell concentrations within which donors can be grouped into responders and non-responders. More significantly, this division seems to be largely independent of the particular tumour cell line examined. The tumour cell lines used here show different

0028–0836/81/260661—02$01.00

Table 1 Aggregation responses of individual donors to six human tumour cell lines

	Cell line U87MG (10^5 ml^{-1})		Hut 20 (10^5 ml^{-1})		HT 29 (10^6 ml^{-1})		Hut 28 (5×10^6 ml^{-1})		SKNMC (5×10^6 ml^{-1})		A549 (10^7 ml^{-1})	
Donor	%	*T*	%	*T*	%	*T*	%	*T*	%	*T*	%	*T*
1	100	2.7	100	2.8	100	2.2	50	4.0	60	3.0	0	—
2	100	2.2	100	2.6	100	3.0	50	4.5	60	3.2	0	—
3	100	3.1	0	—	0	—	0	—	0	—	0	—
4	100	3.5	0	—	0	—	0	—	0	—	0	—
5	100	2.6	90	6.9	0	—	0	—	0	—	0	—
6	100	2.0	100	2.0	100	4.1	50	3.5	60	2.7	0	—
7	40	4.2	0	—	0	—	0	—	0	—	0	—
8	100	2.2	100	3.0	100	2.5	50	3.2	60	3.0	0	—
9	100	3.0	100	4.4	0	—	0	—	0	—	0	—
10	100	3.2	100	3.0	0	—	0	—	0	—	0	—
11	100	2.8	100	3.0	100	3.0	50	3.0	60	2.5	0	—
12	100	2.3	100	2.9	100	3.0	50	3.8	60	3.0	0	—
13	100	2.9	100	3.0	100	3.5	50	4.2	60	3.5	0	—
14	100	3.5	0	—	0	—	0	—	0	—	0	—
15	100	2.6	0	—	0	—	0	—	0	—	0	—
16	100	3.5	0	—	0	—	0	—	0	—	0	—

% Aggregation responses compared with 10μM ADP. *T*, lag time (min) from addition of tumour cells to onset of aggregation.

aggregation profiles and different effects are observed after treatment with apyrase, hirudin and phospholipases[9], which suggests that several different mechanisms may be involved. However, the responses of individual donors to the various cell types observed here are remarkably constant. Of the four cell types in which different responses were observed (Hut 20, HT 29, Hut 28 and SKNMC), six of the donors (3, 4, 7, 14, 15 and 16) showed no response with any of the lines while two other donors (5 and 9) showed a response only with Hut 20 but with a prolonged lag time to the onset of aggregation. Only in the case of one donor (10) did there seem to be normal aggregation with Hut 20 and no aggregation with the other three lines. Further studies are necessary to determine whether a graduated response can be shown at low concentrations of U87MG cells although aggregation by this line seems to involve a unique mechanism.

These observations suggest a greater sensitivity of certain individual donors to platelet aggregation by human tumour cells in general, which may explain why certain types of tumours can be metastatic in one individual but not in another[1]. The results may have more general significance in their application to the epidemiology and genetics of human neoplastic disease.

This work was supported in part by USPHS grants HL 14697, HL 20971 and Biomedical Research Support grant RR 05737. We thank Dr Jorgen Fogh for providing us with the U87MG, HT29 and SKNMC lines, and Dr Addi Gazdar for the Hut 20, Hut 28 and A549 lines.

Note added in proof: Dr Gazdar has since informed us that the Hut 20 line has now been identified as being of murine origin, based on cytogenetic analysis and isozyme phenotyping, while a human origin has been confirmed for the other tumour cell lines examined here. This removes discrepancies associated with the greater reactivity of Hut 20 cells with platelets and explains why 9 of the 16 donors (3–5, 7, 9, 10 and 14–16) failed to respond at the critical tumour cell concentration with all of the differentially reactive human lines HT 29, Hut 28 and SKNMC.

Received 19 November 1980; accepted 21 April 1981.

1. Sugarbaker, E. V. & Ketcham, A. S. *Semin. Oncol.* **4,** 19–32 (1977).
2. Baumgartner, H. R. *Microvasc. Res.* **5,** 167–179 (1973).
3. Marcum, J. M., McGill, M., Bastida, E., Ordinas, A. S. & Jamieson, G. A. *J. Lab. clin. Med.* **96,** 1046–1053 (1980).
4. Gasic, G. J., Kock, P. A. G., Hsu, B., Gasic, T. B. & Niewiarowski, S. *Z. Krebsforsch.* **86,** 263–277 (1976).
5. Holme, R., Oftebro, R. & Hovig, T. *Thromb. Haemostasis* **40,** 89–102 (1978).
6. Pearlstein, E., Cooper, L. B. & Karpatkin, S. *J. Lab. clin. Med.* **93,** 332–344 (1979).
7. Hara, Y., Steiner, S. & Baldini, M. G. *Cancer Res.* **40,** 1212–1216 (1980).
8. Gastpar, H. *Thromb. Diath. haemorrh.* **42,** 291–303 (1970).
9. Bastida, E., Ordinas, A. & Jamieson, G. A. *Cancer Res.* (submitted).

Routes of conjugation in normal and cancerous tissue from human lung

Gerald M. Cohen*, Elizabeth M. Gibby & Rekha Mehta

Toxicology Unit, Department of Pharmacology, School of Pharmacy,
University of London, 29/39 Brunswick Square,
London WC1N 1AX, UK
Department of Biochemistry, University of Surrey,
Guildford GU2 5XH, Surrey, UK

The selective toxicity of drugs leading to major advances in antibacterial chemotherapy has often resulted from the identification and exploitation of major biochemical differences between bacterial and mammalian species[1]. Similar progress has not been made in cancer chemotherapy, partly due to a lack of suitable biochemical differences between normal and cancerous tissue other than in DNA synthesis, but also because of many other problems such as those of metastases and resistance, and the presence in tumours of cells at different states of the cell cycle. Here we report a major biochemical difference in the routes of conjugation between normal lung and tumour tissue from patients with lung cancer. Conjugation with glucuronic acid and sulphate constitute two of the most important pathways of metabolism of drugs, other foreign compounds and hormonal steroids[2,3]. Using 1-naphthol as a model phenolic substrate, normal peripheral lung tissue formed almost exclusively the sulphate ester conjugate, 1-naphthyl sulphate, whereas tumour tissue from squamous carcinomas from the same patients formed predominantly the glucuronic acid conjugate, 1-naphthyl glucuronide. Such major biochemical differences may be exploitable in the design of selectively toxic cancer chemotherapeutic agents.

Peripheral lung tissue and tumours from patients with lung cancer were cultured within 3 h of surgical removal. When short-term organ cultures of peripheral human lung were incubated with [1-^{14}C]1-naphthol, the major metabolite formed was the sulphate conjugate, 1-naphthyl sulphate, with little or no glucuronic acid conjugate being formed (Table 1), which agrees with our previous results[4]. However, in the case of tumour tissue, the metabolism of 1-naphthol varied with the histological classification of the tumour (Table 1). Particularly striking was the finding that tumour tissue from all patients with

---

* To whom correspondence should be addressed.

0028-0836/81/260662—03$01.00

Table 1 Conjugation of [1-^{14}C]1-naphthol by short-term organ cultures of normal human lung and tumour tissue

		% Of total radioactivity recovered in medium as conjugates in the:			
		Lung		Tumour	
Patient	Histological diagnosis	Sulphate	Glucuronide	Sulphate	Glucuronide
VII	Poorly differentiated squamous cell	50.2	ND	0.7	39.8
XI	Poorly differentiated squamous cell	23.7	ND	1.0	8.3
X	Squamous cell	18.9	ND	ND	2.9
XXI	Very poorly differentiated squamous cell	11.5	1.0	1.1	6.9
XVI	Moderately well differentiated squamous cell	16.6	ND	0.1	7.7
XII	Undifferentiated large cell, possibly squamous	35.9	0.3	0.7	12.3
XVIII	Undifferentiated large cell, possibly squamous	11.3	2.0	0.7	13.4
XVII	Undifferentiated large cell	11.1	ND	3.4	6.2
XX	Undifferentiated large cell, possibly adenocarcinoma	13.7	4.4	2.0	3.9
IX	Poorly differentiated adenocarcinoma of large cell, undifferentiated	7.1	ND	ND	ND
XIV	Poorly differentiated adenocarcinoma (possibly metastatic from breast)	27.0	3.1	1.8	4.5
XXII	Small (large) cell anaplastic carcinoma	84.4	1.0	28.5	3.0
XXIII	Large (small) cell anaplastic carcinoma	61.0	3.7	22.0	8.1

Short-term organ cultures of macroscopically normal human lung or tumour tissue were cultured for 18 h at 37 °C either in Leibovitz L-15 medium as previously described[4] or on gelatin sponge in 2.5 ml of CMRL-1066 medium (Biocult, Scotland) as described by Stoner *et al.*[17]. After 18 h the culture medium was changed for one containing [1-^{14}C]1-naphthol (10–20 μM) (specific activity 19.2 Ci mmol^{-1}; Radiochemical Centre) and generally incubated for a further 90 min, but sometimes for another 24 h (patients XXII and XXIII). The amount of radioactivity in the medium at the end of the culture was 70–90% of the total. Aliquots of the media were analysed for 1-naphthol conjugates by TLC as described elsewhere[4]. The results are generally expressed as the mean values from two different explants. Considerable variability was occasionally observed, particularly for the tumour tissue probably due to its heterogeneity. However, the results always showed a similar pattern. Similar metabolite profiles were obtained when lung and tumour were cultured with naphthol: 100–1,000 μM or 5–100 μM respectively. ND, not detectable.

squamous cell carcinomas metabolized 1-naphthol predominantly to its glucuronic acid conjugate, with little or no sulphate conjugate being formed (Table 1). Limited results from other types of lung tumours prevent detailed analysis but show a greater variability in the routes of conjugation (Table 1). These results are supported by our recent studies with human lung cancer cell lines, when cultures derived from squamous cell carcinomas showed extensive formation of 1-naphthyl glucuronide with little sulphate conjugation[5].

To confirm and extend these major differences observed for short-term organ cultures, the rates of conjugation of 1-naphthol by 10,000*g* supernatant fractions of normal lung or tumour tissue were also determined. In patients with sufficient tissue for study, rates of glucuronic acid conjugation were much greater than sulphate ester conjugation in the tumour tissue, whereas in normal lung the reverse was true (Table 2), thus confirming the differences observed using cultured explants.

It is important to stress that the present studies have used lung tissue but most human lung cancers arise from the bronchial epithelium[6]. This should not be a disadvantage as sulphate conjugation is quantitatively far more important than glucuronic acid conjugation for cultured human bronchus[7]. Due to the multiplicity of both UDP-glucuronosyltransferase[2] and sulphotransferase[3] activities, it is difficult to predict the metabolic fate of different substrates from results obtained for one substrate. Nevertheless, with 1-naphthol as substrate, tumorigenesis in human lung is accompanied by the appearance of a conjugation pathway which has little or no activity in the host lung. These changes in conjugation may be affected by alterations in β-glucuronidase and aryl sulphatase, which may be elevated in tumours[8]. However, higher activity of β-glucuronidase in tumours would only minimize the changes noted here. Other tumours may also show similar changes in conjugation; a marked shift from sulphation to glucuronidation was observed in recent studies on cultures of human colon and colonic tumour tissue[9]. The relationship, if any, of these changes in conjugation to tumour development and progression remains obscure, although they may be connected to alterations in cell-surface glycosaminoglycans.

Various rodent hepatomas have shown similar increases in UDP-glucuronosyltransferase activity, accompanied by greatly decreased phenolsulphotransferase activity in a Reuber H-35 hepatoma[10–12]. These results and those reported here showing increased activities of UDP-glucuronosyltransferase in both human and rodent tumours, are in contrast to the generally reported absence or marked decreases in mixed function oxidase activities in both pre-neoplastic nodules and tumours[13,14]. Such alterations in metabolism may have important implications in the successful chemotherapy of neoplastic disease. As some anticancer agents, for example, cyclophosphamide, require metabolism to exert their activity, the balance of oxidative and conjugating enzymes in the tumour will be one of the critical factors determining the amounts of reactive metabolites generated and thus the efficacy of the treatment.

Although conjugation of drugs is generally considered to be a detoxification process, there are examples of conjugation, including glucuronidation[15,16], which lead to more reactive metabolites. Based on the biochemical differences observed

Table 2 Rates of conjugation of [1-^{14}C]1-naphthol by 10,000*g* supernatant fractions of normal human lung and tumour tissue

	Rate of conjugation (nmol per min per g protein) in:			
	Lung		Tumour	
Patient	With sulphate	With glucuronic acid	With sulphate	With glucuronic acid
VII	15.9	0.8	1.9	4.9
XII	11.6	3.4	2.7	16.5
XVII	85.2	3.6	7.5	27.9
XVIII	8.6	3.5	ND	29.1
XXI	6.9	1.7	0.1	100.2

Homogenates (50%) of normal human lung and tumour tissue were centrifuged at 10,000*g* for 30 min and the supernatant fractions either frozen overnight at −20 °C before analysis (for patients VII, XII and XXI) or used as soon as prepared (for patients XVII and XXI). UDP-glucuronosyltransferase activity was measured at 37 °C in 0.5 ml volume containing 2–6 mg protein, 1.5 mM UDP-glucuronic acid (Sigma), 4 mM $MgCl_2$, 20 μM [1-^{14}C]1-naphthol (specific activity 19.2 Ci mol^{-1}) and 0.22 M Tris-HCl buffer, *p*H 7.4. Phenolsulphotransferase activity was measured at 37 °C in 0.5 ml volume containing 3–6 mg protein, 2 mM Na_2SO_4, 5 mM $MgCl_2$, 5 mM ATP (disodium salt, Sigma), 20 μM [1-^{14}C]1-naphthol and 0.14 M Tris-HCl buffer, *p*H 7.4. Control incubations contained no enzyme fractions. Incubations were carried out for 10–60 min to ensure linearity of the reaction. Reactions were stopped by removal of 100-μl aliquots of the incubation media into tubes with an equal volume of methanol containing authentic 1-naphthol and metabolites, which were separated and analysed as described previously. ND, not detectable.

here, it may be possible to design drugs, activated by glucuronic acid conjugation but not by sulphate conjugation, which would be selectively toxic to the tumour cells.

We thank Mr Meredith-Brown and colleagues for the supply of tissues and Professors P. Grasso and B. Corrin for histological diagnosis of the tumours. This work was supported by a grant from the Cancer Research Campaign of Great Britain.

Received 19 January; accepted 1 April 1981.

1. Albert, W. *Selective Toxicity: The Physico-chemical Basis of Therapy* (Chapman & Hall, London, 1973).
2. Dutton, G. J., Wishart, G. J., Leaky, J. E. A. & Goheer, M. A. in *Drug Metabolism—From Microbe to Man* (eds Parke, D. V. & Smith, R. L.) 71–90 (Taylor & Francis, London, 1977).
3. Dodgson, K. S. in *Drug Metabolism—From Microbe to Man* (eds Parke, D. V. & Smith, R. L.) 91–104 (Taylor & Francis, London, 1977).
4. Mehta, R. & Cohen, G. M. *Biochem. Pharmac.* **28,** 2479–2484 (1979).
5. Gibby, E. M., Mehta, R., Ellison, M. & Cohen, G. M. *Biochem. Soc. Trans.* **9,** 111 (1981).
6. Wynder, E. L. & Hecht, S. (ed.) *Lung Cancer, UICC tech. Rep. Ser.* **25,** 3–33 (International Union Against Cancer, Geneva, 1976).
7. Autrup, H. *et al. Int. J. Cancer* **25,** 293–300 (1980).
8. Fishman, W. H. & Anlyan, A. J. *Cancer Res.* **7,** 808–817 (1947).
9. Cohen, G. M., Gibby, E. M., Grafström, R. & Harris, C. C. (in preparation).
10. Lueders, K. K., Dyer, H. M., Thompson, E. B. & Kuff, E. L. *Cancer Res.* **30,** 274–279 (1970).
11. Gessñer, T. *Biochem. Pharmac.* **23,** 1809–1816 (1974).
12. Zimmerman, J. J., Gorski, J. P. & Kasper, C. B. *Drug Metab. Disposit.* **5,** 572–578 (1977).
13. Strobel, H. W., Dignam, T. D., Saine, S. E., Fang, W. & Fennell, P. M. *Molec. cell. Biochem.* **22,** 79–91 (1978).
14. Cameron, R., Sweeney, G. D., Jones, K., Lee, G. & Farber, E. *Cancer Res.* **36,** 3888–3893 (1976).
15. Miller, E. C. *et al. Molec. Pharmac.* **4,** 147–154 (1968).
16. Mulder, G. J., Hinson, J. A. & Gillette, J. R. *Biochem. Pharmac.* **27,** 1641–1649 (1978).
17. Stoner, G. D. *et al. Lab. Invest.* **38,** 685–692 (1978).

Sequential appearance of Epstein–Barr virus nuclear and lymphocyte-detected membrane antigens in B cell transformation

D. J. Moss, A. B. Rickinson, L. E. Wallace & M. A. Epstein

Department of Pathology, University of Bristol Medical School, University Walk, Bristol BS8 1TD, UK

One of the central questions of Epstein-Barr (EB) virus biology concerns the nature of the asymptotic carrier state[1,2] whereby the virus, which *in vitro* transforms B cells into permanent lymphoblastoid cell lines with such efficiency[3,4], is harboured for life in the B lymphoid tissues of all previously-infected (seropositive) individuals[5]. It seems highly probable that this persistent infection is primarily controlled by memory T cells (cytotoxic precursors) specifically primed to the EB virus-induced lymphocyte-detected membrane antigen LYDMA[6–9]. Thus the degree to which the infection can progress without interruption in individual B cells of the immune host will depend upon the timing of LYDMA expression relative to that of other viral gene functions. We now demonstrate that during EB virus-induced B cell transformation *in vitro* LYDMA is expressed on the cell surface soon after the appearance of the virus-associated nuclear antigen EBNA[10] and coincident with, but not dependent upon, the initiation of cellular DNA synthesis. This suggests that in the lymphoid tissues of the infected host, immunological control over virus-B cell interactions can be exercised early in the cycle at the very onset of virus-induced cell proliferation.

EB virus-specific memory T cells are present exclusively in seropositive individuals and can be reactivated *in vitro* by co-cultivation either with freshly-infected autologous B cells[6] or with low numbers of cells from the autologous EB virus-transformed B cell line[11]. Such reactivated T-cell preparations kill virus-infected cells through HLA-A and B antigen-restricted recognition of the target antigen LYDMA[11–15], a virus-induced

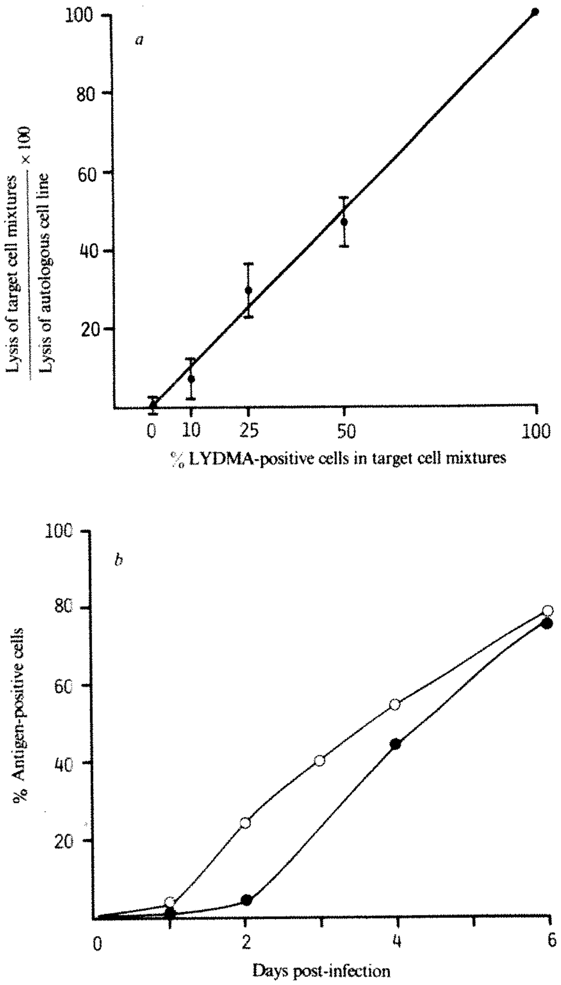


Fig. 1 *a*, Calibration curve of activity of effector T cells from donor D.M. Freshly prepared blood mononuclear cells from donor D.M. were seeded at 2×10^6 cells per 2-ml well (Linbro) with 2.5×10^4 cells of the autologous EB virus-transformed cell line. At 10 days later the E-rosette-forming subpopulation was prepared by centrifugation over Ficoll–Hypaque followed by lysis of the erythrocytes in NH_4Cl. Replicate ampoules of these effector T cells were stored in liquid nitrogen. Full details of these procedures are given elsewhere[9,11]. Effector T cells from one such ampoule were tested in a preliminary chromium release assay for cytotoxicity against the autologous EB virus-transformed cell line (DM-EB), against EB virus-transformed cell lines from three HLA-A and -B antigen-unrelated donors M.O., J.W. and A.S. (MO-EB, JW-EB, AS-EB) and against two EB virus genome-negative cell lines K562 and HSB-2. Cytotoxicity was expressed in terms of the per cent specific lysis of the ^{51}Cr-labelled target cells observed following a 5-h incubation with effector T cells in V-shaped wells at an effector:target ratio of 20:1, as described elsewhere[13]. Levels of cytotoxicity were as follows: DM-EB, 52%; MO-EB, 8%; JW/EB, 2%; AS-EB, 3%; K562, 2%; HSB-2, 6%. Effector T cells from other ampoules of this preparation were then tested for cytotoxicity against autologous T-depleted (TD) mononuclear cells cultured without exposure to EB virus, against the autologous EB virus-transformed cell line and against artificial mixtures containing these two different types of target cells in varying proportions. The calibration curve shown was constructed by combining the results of three separate tests of this kind, carried out on days 1, 2 and 4 of an experiment with cells from donor D.M. For each artificial target cell mixture, the extent of lysis was on each occasion expressed as a percentage of that simultaneously recorded against the autologous target cell line itself and the mean percentage (±s.d) from the three tests was then plotted against the proportion of LYDMA-positive cells (that is, cells of the autologous line) in that mixture. The absolute level of autologous cell line lysis recorded on these occasions was 55–70% specific isotope release; the absolute level of lysis of cultured autologous TD cells was always below 5%. *b*, Induction of EBNA-positive (○) and LYDMA-positive (●) cells in cultures of EB virus-infected TD cells from donor D.M. The percentage of EBNA-positive cells in each test population was obtained from counts on 1,000 cells. All appropriate controls were included within each EBNA-staining procedure, as described elsewhere[29]. Only cells showing bright unequivocal staining with positive serum were scored as EBNA positive; preparations stained with negative serum showed virtually no cells with nuclear fluorescence. The percentage of LYDMA-positive cells in each test population was deduced from the extent of specific lysis in the chromium release assay (relative to the specific lysis of the autologous cell line itself) using the calibration curve of effector T-cell activity shown in *a*.

0028-0836/81/260664—03$01.00

cell surface change whose existence was first invoked to explain the cellular response to primary EB virus infection seen in infectious mononucleosis patients[16–18]. Thus strong killing of the autologous, and of allogeneic HLA-A and B antigen-related, EB virus-transformed cell lines by *in vitro* reactivated T cells occurred in the absence of any significant lysis of (1) the corresponding cell lines from HLA-unrelated donors, (2) autologous pokeweed mitogen (PWM)-stimulated B lymphoblasts, or (3) EB virus genome-negative human haemopoietic cell lines such as K562 and HSB2 deliberately included as sensitive indicators of non-specific or natural killer-like activities which under other circumstances[19–22] can arise in cultured T-cell populations (see Fig. 1 legend, Fig. 2 and refs 11, 13, 15).

Blood mononuclear cells from seropositive donors were co-cultivated with cells of the autologous EB virus-transformed cell line at a ratio of 80 : 1 and the E-rosette-forming subpopulation collected from these co-cultures 8 to 12 days later provided EB virus-specific cytotoxic T cell preparations which were stored in replicate aliquots in liquid nitrogen until required[23]. In the present experiments these effector T cells have been used to monitor the appearance of LYDMA during the virus-induced *in vitro* transformation of autologous B cells. Freshly prepared blood mononuclear cells from the relevant donors were first depleted of T cells by E-rosetting and one part of the T-depleted (TD) subpopulation was then exposed to a potent EB virus preparation, and the other left uninfected, before seeding in Linbro plates at 1.5×10^6 cells per 2-ml well. On at least two occasions during each experiment, the relevant effector T cells were tested in a 5-h chromium release assay against autologous TD cells from the uninfected cultures, against mixtures containing these TD cells and a 10, 25 or 50% proportion of cells from the autologous virus-transformed line, and against the autologous cell line itself. The results provided a calibration curve of effector T-cell activity, such as that shown in Fig. 1*a* for effectors from seropositive donor D.M., and confirmed that the assay was sufficiently sensitive to detect as few as 10% virus-transformed (LYDMA-positive) cells in a mixed target cell population.

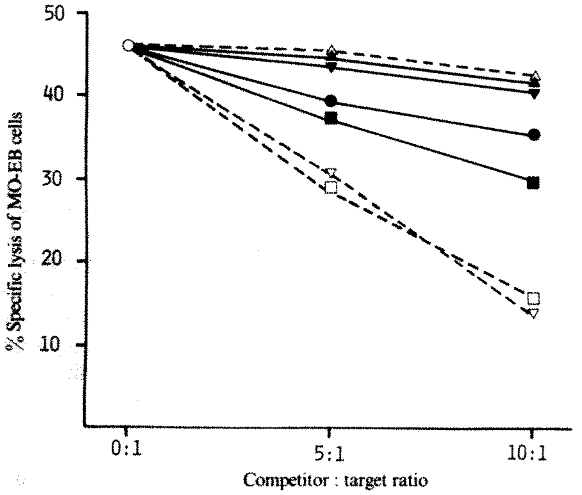


Fig. 2 Induction of LYDMA-positive cells in cultures of EB virus-infected TD cells from donor M.O. assayed by their capacity to serve as competitors of EB virus-specific cytotoxicity. Cultures of EB virus-infected TD cells from donor M.O. were collected and the cells stored in liquid nitrogen 1, 3, 4 and 7 days postinfection; in these preparations immunofluorescence staining showed here to be 1, 33, 53 and 74% EBNA-positive cells respectively. Effector T cells were prepared from donor S.G. (matched to donor M.O. at all four HLA-A and -B loci) and were tested for cytotoxicity against ^{51}Cr-labelled target cells of the EB virus-transformed cell line from donor M.O. (MO-EB) either alone (○) or in the presence of the following unlabelled competitor cells: EB virus-infected MO-TD cells collected 1 (▲), 3 (▼), 4 (●) and 7 (■) days post-infection, prepared as above and thawed immediately before use; EB virus-transformed MO-EB cells (□); EB virus-transformed SG-EB cells (▽); TD cells from donor S.G. stimulated for 6 days in the presence of PWM and autologous mitomycin C-treated T cells as described elsewhere[13] (△). When simultaneously tested against a range of target cells, the effector T cells from donor S.G. used above produced the following levels of specific lysis on (1) EB virus-transformed target cells: SG-EB (autologous), 42%; MO-EB (HLA-A and -B antigen-identical), 46%; HD-EB (HLA-A and -B antigen-unrelated), 3%; JR-EB (HLA-A and -B antigen-unrelated), 1%; (2) EB virus genome-negative target cells: PWM-stimulated SG-TD cells (autologous), −7%; PWM-stimulated MO-TD cells (HLA-A and -B antigen identical), −3%; K562, 1%; HSB-2, 7%.

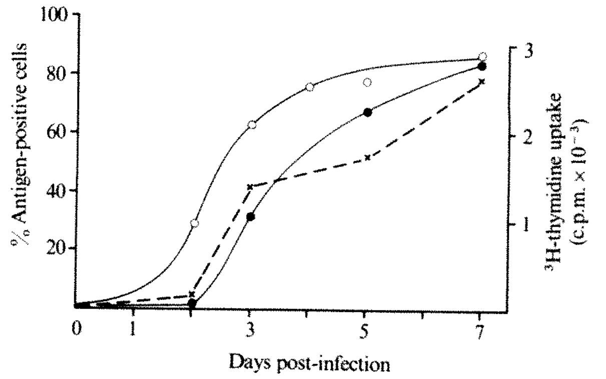


Fig. 3 Induction of EBNA-positive cells (○), LYDMA-positive cells (●) and DNA synthesis (×) in cultures of EB virus-infected TD cells from donor S.G. Assays for EBNA and LYDMA were as described in Fig. 1 legend. The effector T cells from donor S.G. used to detect LYDMA-positive cells in this experiment showed a pattern of EB virus-specific, HLA-A and -B antigen-restricted cytotoxicity similar to that described in Fig. 2 legend; the absolute level of autologous cell line lysis mediated by these effector T cells was 50–60% specific chromium release. Cellular DNA synthesis was measured by labelling the cultures with ^{3}H-thymidine (TRA61, Amersham; specific activity 5 Ci mM^{-1}; concentration 0.5 μCi ml^{-1} culture medium) for 18 h immediately before collection. Replicate 20-μl samples from each cell suspension were washed well with saline, followed by 10% ice-cold trichloroacetic acid and ice-cold methanol. The filters were dried and the remaining radioactivity measured by scintillation counting. The difference between the mean count obtained from virus-infected cultures and that obtained from parallel uninfected cultures was taken to indicate the level of virus-induced DNA synthesis.

Virus-infected TD cell cultures were collected on up to five occasions during the first week post-infection (including those days on which the above calibrations were performed), the cells counted and then used either to prepare slides for immunofluorescence staining to determine the proportion of EBNA-positive cells, or as targets for the relevant effector T cells in a chromium release assay (using the autologous cell line as a standard target on each occasion) to monitor the appearance of LYDMA-positive cells. All seven donors tested in this initial series of experiments gave results of the kind shown in Fig. 1*b*, again from donor D.M. Bright EBNA-positive staining, detectable in a small proportion of cells on day 1, was clearly present in 25% or more cells by day 2, rising steadily thereafter to levels of around 80% within 6–7 days. By comparison, virus-infected TD cells never suffered significant lysis in chromium release assays on days 1 or 2 post-infection but thereafter became increasingly sensitive such that about 80% of the cells were detectably LYDMA-positive by days 6–7.

The above chromium release data seems to give a true indication of the kinetics with which cells develop the T-lymphocyte-detected antigen (rather than merely reflecting some temporal change in the sensitivity of antigen-bearing cells to lysis *per se*) because a very similar pattern of results was obtained on three further occasions when the appearance of LYDMA-positive cells was monitored by testing virus-infected TD cell populations as unlabelled competitors of EB virus-specific cytotoxicity. Figure 2 shows the results from one such experiment using the EB virus-transformed cell line from donor M.O. as the source of target cells and a reactivated T-cell preparation from the HLA-A and -B antigen-identical donor S.G. as the source of effectors. When sequentially collected populations of EB virus-infected MO-TD cells were included as competitors in this assay, they did not begin to cause significant inhibition of lysis (that is, above the very low background effect of PWM-stimulated B-lymphoblast competitors) until day 4 post-infection even though some 33% EBNA-positive cells were already detectable by day 3. The discordance between the two antigen levels which was consistently observed on days 2–3 post-infection, and the rapid increase to equivalent values thereafter, strongly suggest that the appearance of EBNA in the nuclei of

Table 1 Development of EBNA and LYDMA in thymidine-blocked cultures

Cells in culture	Treatment	Day of collection	Cell count per ml	% EBNA-positive cells	% LYDMA-positive cells
SG-TD	EBV	5	0.35×10^6	78	68
		7	10^6	86	84
SG-TD	EBV + 10 mM Tdr	5	0.12×10^6	70	81
		7	0.15×10^6	76	88

Unlabelled thymidine (Tdr) was added to the cultures 24 h post-infection to a final concentration of 10 mM. The establishment of a block to cell proliferation, exerted at the first G_1-S boundary in thymidine-treated cultures, was confirmed by Hoechst 33342-staining analysis[24] of the cell populations shown above; when analysed in a fluorescence activated cell sorter, >95% thymidine-treated cells had pre-synthetic DNA values characteristic of the G_1 phase, whereas non-thymidine-treated cells showed a heterogeneous spread of DNA values across the cell cycle (data not shown).

virus-infected B cells precedes by about 24 h the expression of LYDMA in a recognizable form on the cell surface.

In three of the seven experiments of the initial series, both control and virus-infected TD cell cultures were labelled with ^{3}H-thymidine for 18 h immediately before collection so as to assess the degree of virus-induced cell proliferation, in addition to the levels of viral antigen expression. Typical results, with cultures from donor S.G. (Fig. 3), show the induction of LYDMA and the initiation of cellular DNA synthesis to be coincident events occurring soon after the appearance of EBNA in virus-infected cells. In these same experiments, exposure of some virus-infected cultures to 10 mM unlabelled thymidine from day 1 onwards stopped the subsequent entry of cells into S phase, a block confirmed by Hoechst 33342-staining analysis of the cell populations in a fluorescence-activated cell sorter[24], but did not prevent a substantial majority of the infected cells going on to develop both EBNA and LYDMA (Table 1).

A close temporal relationship between the expression of EBNA, a DNA-binding protein thought to induce cell proliferation[25], and that of LYDMA, a cell-surface change which elicits T-cell-mediated attack[15,18], has important implications for the biology of the EB virus carrier state. Two facts, the remarkably strong long-lived T-cell surveillance which EB virus infection usually induces in its host[6], and the obligatory appearance of LYDMA on cells at an early stage of the infectious cycle (Figs 1*b*, 2) coincident with, but not dependent on, the initiation of cell proliferation (Fig. 3, Table 1), are difficult to accommodate within that current view of EB virus persistence which postulates the long-term survival of individual virus-transformed B cells in lymphoid tissues. This raises the possibility of an alternative view, namely that the virus might persist by way of chronic, usually low grade, replication within the oropharynx[26–28], involving some specialized cell and/or some immunologically privileged site, and that the apparent virus-carrier status of lymphoid tissues may be maintained only by the continual in-flow of newly infected B cells from that site coupled with their continual removal by immune T cells once the stage of LYDMA expression has been reached. Relaxation of T-cell surveillance in immunologically compromised individuals would of course allow EBNA/LYDMA-positive cells to begin proliferating and this may be a necessary first step in the pathogenesis of EB virus-associated malignant disease.

We thank Elizabeth Gough and Ann Green for technical help, and Dr B. Thomas and Mr K. Keeler for advice and assistance with the fluorescence-activated cell sorter analysis of Hoechst 33342-stained cell preparations. This work was supported by MRC grants G976/412 and G979/1233, and was assisted by the Bradbury Investment Company of Hong Kong. D.J.M. was supported by a D.B. Duncan Cancer Research Fellowship while on leave from the Queensland Institute of Medical Research, Herston, Brisbane, Australia.

Received 5 November 1980; accepted 17 March 1981.

1. Epstein, M. A. & Achong, B. G. *A. Rev. Microbiol.* **31,** 421–445 (1977).
2. Klein, G. *Cancer* **45,** 2486–2499 (1980).
3. Henle, W., Diehl, V., Kohn, G., zur Hausen, H. & Henle, G. *Science* **157,** 1064–1065 (1967).
4. Pope, J. H., Horne, M. K. & Scott, W. *Int. J. Cancer* **3,** 857–866 (1968).
5. Nilsson, K., Klein, G., Henle, W. & Henle, G. *Int. J. Cancer* **8,** 443–450 (1971).
6. Moss, D. J., Rickinson, A. B. & Pope, J. H. *Int. J. Cancer* **22,** 662–668 (1978).
7. Rickinson, A. B., Moss, D. J. & Pope, J. H. *Int. J. Cancer* **23,** 610–617 (1979).
8. Moss, D. J., Rickinson, A. B. & Pope, J. H. *Int. J. Cancer* **23,** 618–625 (1979).
9. Rickinson, A. B., Wallace, L. E. & Epstein M. A. *Nature* **283,** 865–867 (1980).
10. Reedman, B. M. & Klein, G. *Int. J. Cancer* **11,** 499–520 (1973).
11. Rickinson, A. B., Moss, D. J., Allen, D. J., Wallace, L. E., Rowe, M. & Epstein, M. A. *Int. J. Cancer* (in the press).
12. Misko, I. S., Moss, D. J. & Pope, J. H. *Proc. natn. Acad. Sci. U.S.A.* **77,** 4247–4250 (1980).
13. Moss, D. J., Wallace, L. E., Rickinson, A. B. & Epstein, M. A. *Eur. J. Immun.* (in the press).
14. Wallace, L. E., Moss, D. J., Rickinson, A. B., McMichael, A. J. & Epstein, M. A. *Eur. J. Immun.* (in the press).
15. Rickinson, A. B., Moss, D. J., Wallace, L. E., Rowe, M., Misko, I. S., Epstein, M. A. & Pope, J. H. *Cancer Res.* (in the press).
16. Svedmyr, E. & Jondal, M. *Proc. natn. Acad. Sci. U.S.A.* **72,** 1622–1626 (1975).
17. Royston, I., Sullivan, J. L., Periman, P. O. & Perlin, E. *New Engl. J. Med.* **293,** 1159–1163 (1975).
18. Klein, E., Klein, G. & Levine, P. H. *Cancer Res.* **36,** 724–727 (1976).
19. Ortaldo, J. R., Bonnard, G. D., Kind, P. D. & Herberman, R. B. *J. Immun.* **122,** 1489–1494 (1979).
20. Seeley, J. K., Masucci, G., Poros, A., Klein, E. & Golub, S. H. *J. Immun.* **123,** 1303–1311 (1979).
21. Viallat, J., Svedmyr, E., Steinitz, M. & Klein, G. *Cell. Immun.* **38,** 68–75 (1978).
22. Jondal, M. & Targan, S. *J. exp. Med.* **147,** 1621–1636 (1978).
23. Van Lambalgen, R., Farrant, J. & Bradley, B. A. *J. immun. Meth.* **27,** 327–338 (1979).
24. Lydon, M. J., Keeler, K. D. & Thomas, D. B. *J. cell. Physiol.* **102,** 175–181 (1980).
25. Luka, J., Siegert, W. & Klein, G. *J. Virol.* **22,** 1–8 (1977).
26. Gerber, P., Nonoyama, M., Lucas, S., Perlin, E. & Goldstein, L. I. *Lancet* **ii,** 988–989 (1972).
27. Miller, G., Niederman, J. C. & Andrews, L-L. *New Engl. J. Med.* **288,** 229–232 (1973).
28. Morgan, D. G., Niederman, J. C., Miller, G., Smith, H. W. & Dowaliby, J. M. *Lancet* **ii,** 1154–1157 (1979).
29. Crawford, D. H., Rickinson, A. B., Finerty, S. & Epstein, M. A. *J. gen. Virol.* **38,** 449–460 (1978).

Inserted H-2 gene membrane products mediate immune response phenotype of antigen-presenting cell

Aya Jakobovits*, Ayala Frenkel†, Nathan Sharon* & Irun R. Cohen†

Departments of Biophysics* and Cell Biology†, The Weizmann Institute of Science, Rehovot, Israel

Soluble protein antigens do not stimulate helper or effector T lymphocytes unless they have first been processed and exposed to the T lymphocytes by appropriate antigen-presenting cells. This antigen presentation was found to be controlled by products of genes in the major histocompatibility complex (MHC), designated H-2 in the mouse. Efficient antigen presentation to T lymphocytes is obtained only when both presenting and responding cells share alleles within the I region of the MHC[1]. Furthermore, immune response (Ir) genes in the MHC influence the choice of antigenic determinants recognized on the immunogenic molecule[2]. To investigate the molecular mechanisms underlying Ir gene expression, we have inserted membrane components isolated from cells of one Ir genotype into recipient cells of a different Ir genotype, using membrane vesicles composed of donor plasma membranes and envelope glycoproteins from Sendai virus[3]. We now report that this 'membrane-engineering' enables genetically inappropriate antigen-presenting cells to communicate to T lymphocytes antigenic determinants of beef or sheep insulins, antigens under control of Ir genes in mice[4]. This demonstrates that structural products of MHC genes, expressed and active at the plasma membrane, function in the Ir phenotype of the response to insulin.

We immunized H-2 congeneic mice, C3H · DiSn (H-2^k) or C3H ·SW (H-2^b), against sheep or beef insulin, respectively. Although these two insulins differ by substitution of only one amino acid at position 9 in the A chain (sheep A9 Gly, beef A9 Ser), the products of H-2-linked Ir genes are sensitive to this structurally conservative difference. Mice of the H-2^b genotype respond well to beef insulin but only poorly to sheep insulin,

0028-0836/81/260666—03$01.00

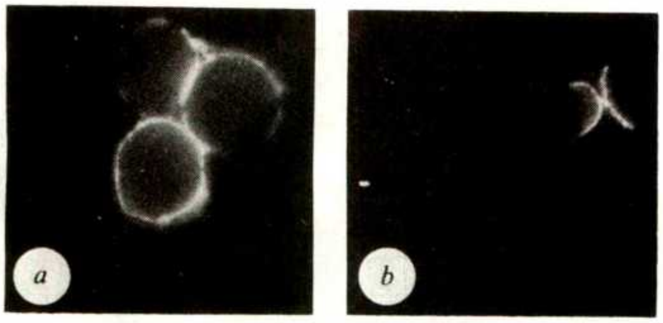


Fig. 1 H–2^{b} antigens inserted into membranes of H–2^{k} acceptor cells are capped by anti-H–2^{b} antiserum, detected by immunofluorescence. C3H · DiSn (H–2^{k}) spleen cells were fused with membrane vesicles containing C3H · SW (H–2^{k}) plasma membranes. Fusogenic membrane vesicles were prepared by mixing 1 mg protein of freshly solubilized plasma membranes with 1 mg protein of solubilized Sendai virus envelopes and dialysing the mixture for 96 h at 4 °C in the presence of Bio-beads SM-2 for efficient removal of the detergent[3]. The dialysed solution was centrifuged (100,000*g*, 3 h, 4 °C) to obtain reconstituted vesicles. Fusion procedure: 10^7 C3H · DiSn (H–2^{k}) spleen cells were suspended in 1 ml of fusion solution (0.14 M NaCl, 20 mM Tris-HCL *p*H 7.4, 3 mM KCL, 0.8 mM $MgSO_4$) and incubated with 5 μg protein of reconstituted membrane vesicles for 60 min at 4 °C, with occasional shaking, during which the vesicles bound to the cells. To complete the process of fusion, the cells were collected by centrifugation, resuspended in 1 ml of fusion solution containing 5 mM Ca^{2+}, incubated at 37 °C for 30 min and then washed twice with phosphate-buffered saline, *p*H 7.4. To detect H–2^{b} antigens on the surface membranes of the H–2^{k} cells (*a*), the cells were incubated for 30 min at 4 °C with anti-H–2^{b} antiserum (prepared by immunizing C3H · DiSn (H–2^{k}) mice with C3H · SW (H–2^{b}) spleen cells) followed by incubation with fluorescein isothiocyanate-labelled protein A for 30 min at 4 °C. To produce capping (*b*), the cells from *a* were incubated for additional 15 min at 37 °C (ref. 7). Negative results were obtained using control antiserum or H–2^{k} acceptor cells that had not been fused with membrane vesicles.

whereas H–2^{k} mice respond only to sheep insulin [4,5]. We examined whether memory T lymphocytes from immunized mice would respond in a secondary proliferation assay *in vitro* to antigen-presenting spleen cells that had been pulsed with one of these insulins. It has been shown that insulin-pulsed spleen cells of an incompatible H–2 genotype will not present insulin to memory T lymphocytes[1], as T lymphocytes seem to recognize their specific antigen only in association with appropriate H–2 gene products[6]. Our approach was to insert membrane components isolated from cells of H–2^{b} or H–2^{k} responder genotype into the membrane of presenting cells of H–2 nonresponder genotype, and to examine whether the inserted H–2 structural products could enable genetically inappropriate cells to present insulin to primed memory T lymphocytes.

To insert donor plasma membrane components into the membrane of recipient cells, we reconstituted isolated plasma membranes of spleen cells together with Sendai virus envelopes to obtain fusogenic membrane vesicles that could be integrated into the membranes of intact acceptor cells without killing them[3]. Isolation of purified plasma membranes and preparation of fusogenic membrane vesicles were carried out as previously described[3]. Antigen-presenting cells were fused with fusogenic membrane vesicles from H–2 congeneic sources and the integration of the new H–2 antigens in the cell membrane was detected by binding specific anti-H–2 antiserum. As shown in Fig. 1, fusion of C3H · DiSn (H–2^{k}) cells with vesicles prepared from C3H · SW (H–2^{b}) spleen cell plasma membranes led to the appearance of H–2^{b} antigens on the surface of H–2^{k} cells. Capping[7] of the inserted H–2 antigens indicated that these antigens were functionally integrated into the membranes of the acceptor spleen cells. Moreover, it had been previously shown that inserted allogeneic H–2 antigens, but not unfused allogeneic vesicles, could also stimulate a primary mixed-lymphocyte reaction in syngeneic responder lymphocytes[3].

Figure 2 shows the results of inserting H–2^{k} or H–2^{b} plasma membranes into H–2 syngeneic or allogeneic antigen-presenting spleen cells. In detecting the specific response to insulin, we had to take into account the fact that allogeneic H–2 antigen endogenous or inserted into spleen cells will stimulate T lymphocytes[8]. The net response to insulin was therefore the c.p.m. obtained when the antigen-presenting spleen cells were pulsed with insulin minus the c.p.m. generated in the absence of insulin pulsing—the stimulation due to allogeneic H–2 antigens was considered as background.

In the first experiment, lymph node lymphocytes were taken from C3H · DiSn (H–2^{k}) responder mice that had been immunized *in vivo* against sheep insulin. Antigen-presenting cells originated from C3H · SW (H–2^{b}) mice which are genetical low responders to sheep insulin. These H–2^{b} spleen cells, either untreated or fused with vesicles containing virus envelope components alone, or together with C3H · SW (H–2^{b}) membrane antigens, stimulated a response to their allogeneic H–2^{b} antigens, but did not stimulate a net response to insulin. However, insertion of C3H · DiSn (H–2^{k}) membrane components into the H–2^{b} antigen-presenting cells led to a strong net response to sheep insulin of 65,000 c.p.m. above the allogeneic background level measured in the absence of insulin. Similarly, fusion of C3H · SW (H–2^{b}) antigen-presenting cells with C3H/eB (H–2^{k}) membranes led to a strong net proliferative response to sheep insulin (62,000 c.p.m.).

The effect of membrane insertion on the ability of spleen cells to present beef insulin was also studied. The responding cells were lymph node cells taken from C3H · SW (H–2^{b}) responder mice that had been primed against beef insulin while antigen-presenting cells originated from C3H · SW (H–2^{b}) responder mice or from C3H · DiSn (H–2^{k}) nonresponder mice. C3H · SW (H–2^{b}) spleen cells presented beef insulin to syngeneic responding cells (net response 28,000 c.p.m.). The insertion of nonresponder allogeneic C3H · DiSn (H–2^{k}) membranes into the C3H · SW (H–2^{b}) spleen cells led to a significant allogeneic stimulation, but the net response to beef insulin was 19,000 c.p.m. about 30% less than that elicited by untreated C3H · SW (H–2^{b}) spleen cells. Antigen-presenting spleen cells from C3H · DiSn (H–2^{k}) mice produced an allogeneic response of about 40,000 c.p.m. with a negligible net response to beef insulin (7,000 c.p.m.). However, fusion of these H–2^{k} spleen cells with C3H · SW (H–2^{b}) membrane vesicles produced a strong net response to beef insulin of 62,000 c.p.m. above the allogeneic background stimulation. In contrast, insertion of vesicles containing virus envelope, either alone or with C3H · DiSn (H–2^{k}) membranes (not shown), into the C3H · DiSn (H–2^{k}) spleen cells did not improve the capacity of these cells to present beef insulin.

These results show that genetically inappropriate antigen-presenting spleen cells can be made to present insulin by inserting membrane components taken from H–2 genetically appropriate spleen cells. This was probably due to specific products of H–2 genes, because the donor and acceptor cells were congeneic and differed only at their H–2 alleles. Presentation of insulin could not be attributed to any nonspecific allogeneic effects[9] because, as reported by others[1], we observed that unmodified allogeneic spleen cells did not present antigens to primed memory T lymphocytes (Fig. 2 *a*, *b*), and insertion of membranes from nonresponder H–2^{k} genotype into H–2^{b} responder spleen cells led to a strong allogeneic response that not only did not enhance, but decreased the presentation of beef insulin by these cells to syngeneic H–2^{b} lymphocytes (Fig. 2*a*). In addition, insertion of viral envelope components alone or together with nonresponder H–2 membranes showed no effect on presentation of insulin. Thus, the ability of the inserted membrane components to change the presentation properties of the recipient spleen cells seemed to be specific to the H–2 alleles

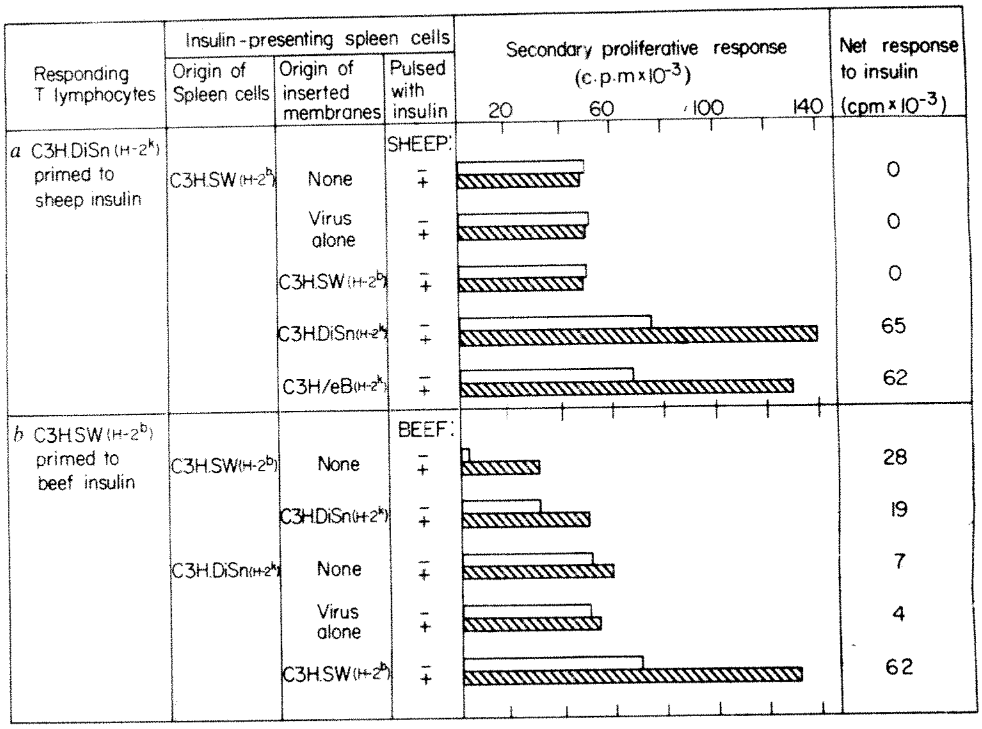


Fig. 2 Inserted H-2 gene products mediate Ir phenotype of spleen cells presenting insulin to primed T lymphocytes. Mice of congeneic strains C3H·DiSn (H-2^k) or C3H·SW (H-2^b) were immunized by injecting 5 μg of insulin emulsified in complete Freund's adjuvant into each hind footpad, as described elsewhere[5]. Nine days later, suspensions of lymphocytes from the draining popliteal lymph nodes were suspended in RPMI-1640 medium supplemented with 0.5% normal mouse serum, HEPES buffer (10×10^{-3} M, *p*H 7.3), non-essential amino acids (Biolab, 1%), sodium pyruvate (1×10^{-3} M), L-glutamine (2×10^{-3} M), 2-mercaptoethanol (5×10^{-5} M) and antibiotics. 4×10^5 cells were seeded in U-shaped microtitre wells (Greiner) in 0.1 ml medium. Antigen-presenting spleen cells were untreated or fused with vesicles as described in Fig. 1 legend. Vesicles containing virus alone were prepared by reconstituting virus envelopes without the addition of plasma membranes. After fusion, the antigen-presenting cells (10^7 per ml RPMI-1640) were incubated with or without insulin (100 μg ml^{-1}) for 60 min at 37 °C, irradiated (1,000 rad) to inhibit their synthesis of DNA and washed twice in medium. The antigen-presenting cells were added (10^5 cells per 0.1 ml medium) to each well containing primed lymph mode lymphocytes and the cells mixtures were incubated for 96 h in quadruplicate cultures. To measure the proliferative response, ^{3}H-thymidine (1 μCi per well) was added to each well for the last 16 h of culture and incorporation into DNA was measured by collecting the cells and measuring the c.p.m. as described[5]. s.e.m. were all <10%. The net response to insulin was computed by subtracting the c.p.m. obtained in cultures without insulin-pulsed antigen-presenting spleen cells from the c.p.m. obtained in cultures with insulin-pulsed antigen-presenting spleen cells. The results of the groups showing net response to insulin of ≥19,000 c.p.m. were highly significant by Student's *t*-test ($P < 0.001$). Non-fusogenic vesicles containing plasma membrane components without viral glycoproteins had no influence on the response to insulin (data not shown).

and Ir genes that distinguished the congeneic strains of mice used.

As well as modifying the Ir phenotype of the insulin-presenting cells insertion of syngeneic H-2 membrane products into allogeneic cells augmented the background allogeneic stimulation (Fig. 2). The recognition of allogeneic antigens may have been enhanced by association with self-H-2 products in the same membrane[6].

It has been suggested that presentation of antigens involves two distinct processes coded by H-2 genes: the communication of signals between cells mediated by 'cell-interaction structures'[10], and the selection of specific antigenic determinants on the immunogenic molecule[2]. Our results indicate that structural products of H-2 genes located at the plasma membrane of the presenting cells are sufficient to carry out these functions. A strategy of membrane reconstruction using isolated and purified H-2 gene products will make it possible to study the molecular mechanisms responsible for these cellular functions.

This study was supported in part by grant AM26766 from the National Institute of Arthritis, Metabolism and Digestive Diseases. Dr P. De Baetselier donated the fluorescein isothiocyanate-labelled Protein A.

Received 8 December 1980; accepted 17 March 1981.

1. Yano, A., Schwartz, R. H. & Paul, W. E. *J. exp. Med.* **146,** 828–843 (1977).
2. Rosenthal, A. S., Barcinski, M. A. & Blake, J. T. *Nature* **267,** 156–158 (1977).
3. Prujansky-Jakobovits, A., Volsky, D. J., Loyter, A. & Sharon, N. *Proc. natn. Acad. Sci. U.S.A.* **77,** 7247–7251 (1980).
4. Keck, K. *Eur. J. Immun.* **7,** 811–816 (1977).
5. Cohen, I. R. & Talmon, J. *Eur. J. Immun.* **10,** 284–289 (1980).
6. Snell, G. D. *Immun. Rev.* **38,** 3–69 (1978).
7. Taylor, R. B., Duffus, W. P. H., Raff, N. C. & De Petris, S. *Nature new Biol.* **233,** 225–229 (1971).
8. Bach, F. H., Widmer, M. B., Bach, M. L. & Klein, J. *J. exp. Med.* **136,** 1430–1444 (1972).
9. Osborne, D. P. Jr & Katz, D. H. *J. exp. Med.* **138,** 825–838 (1973).
10. Katz, D. H. & Benacerraf, B. *Transplantn Rev.* **22,** 175–195 (1975).

Differences between germ-line and rearranged immunoglobulin V_κ coding sequences suggest a localized mutation mechanism

Michael Pech, Josef Höchtl, Hannelore Schnell & Hans G. Zachau

Institut für Physiologische Chemie, Physikalische Biochemie und Zellbiologie der Universität München, 8000 München 2, FRG

Several mechanisms contribute to the generation of an antibody repertoire (for reviews see refs 1–3) which has been estimated to exceed 10^6 different types of immunoglobulin molecules[4]. It is clear that the germ line contains a fair number of V-gene segments[2,5–7]. The V-gene segments are joined by somatic recombination to J- and D-gene segments[8–11] and there is, of course, combinatorial association of heavy and light chains. It has been suggested that the germ-line repertoire of gene segments and their combinatorial joining are the principal sources of antibody diversity[12]. In fact, two rearranged V_κ-gene segments have been found to be identical in sequence to their germ-line counterparts[12,13], showing that no somatic point mutations had occurred. In the genes for a λ_1 light chain[14] and for two heavy chains[10,15], on the other hand, the rearranged V-gene segments differed in several positions from their germ-line equivalents; this is explained by somatic point mutations. We now report the finding of six single-base pair differences each between two rearranged V_κ-gene segments and their presumptive germ-line equivalents and the absence of any such differences in the adjacent non-coding regions; this suggests the existence of localized mutation mechanism(s) in this V_κ gene system.

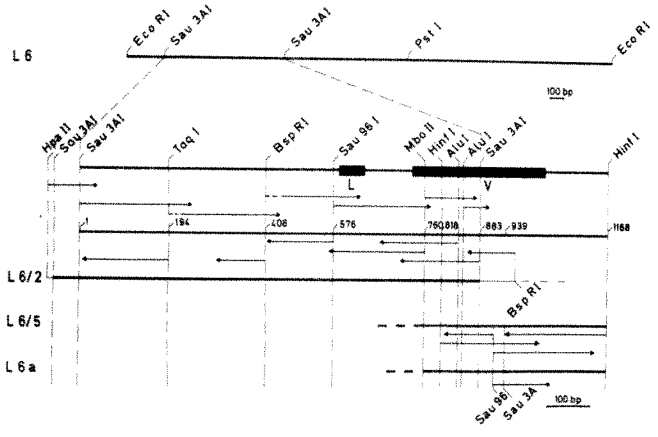


Fig. 1 Partial restriction maps and sequencing strategies of L6 and the subclones L6/2 and L6/5. The subclone L6/2 was constructed by ligation of a *Sau*3AI digest of L6 into the *Bam*HI site of pBR322. For the construction of subclone L6/5 the leftward *Eco*RI–*Pst*I fragment of L6 was ligated with an *Eco*RI–*Pst*I digest of pBR322. In L6/2 the *Sau*3A fragment 884–939 was ligated onto the left end of fragment 1–883. The restriction maps of subclone L6/5 and the second independently isolated λgtWES clone L6a (bottom line) (see text) are shown only in the regions relevant for the sequencing strategy. Thick lines designate mouse DNA, thin lines pBR322 DNA. Digestions, isolation of DNA fragments and further steps of sample preparation were as in ref. 24 with modifications according to ref. 17.

The variable regions of two rearranged κ light-chain genes from the same mouse tumour, myeloma T (ref. 16), have recently been sequenced[17]. The two gene segments which had been generated by joining to the J5 and J2 gene segments (designation as in ref. 8) were called V-T1 and V-T2, respectively[16]. The V-T1 gene segment is expressed (ref. 17 and F. Falkner, unpublished) whereas V-T2 is non-functional because of a one-nucleotide deletion at the V–J junction. Presumptive germ-line counterparts were cloned from liver DNA of BALB/c mice[7,8], which in blotting experiments of digests was identical to embryonic DNA[16]. Column chromatographic fractions of *Eco*RI digests of liver DNA were screened by hybridization with cloned fragments containing V-T1 and V-T2 (Fig. 1 in ref. 7). Five fragments hybridizing with V-T1 were cloned, four of which can be excluded as germ-line equivalents of V-T1 on the basis of restriction analyses[7,18]. Other fragments were excluded because of their weak hybridization signals. The fragment designated L6, however, which gave rise to the strongest hybridization signal, had identical restriction patterns in the V-gene region. The same pattern was found in three independently isolated clones. V-L6 is therefore the presumptive germ-line equivalent of V-T1 (ref. 7). Screening with a V-T2-containing clone revealed three weakly hybridizing fragments which differed from V-T2 in restriction patterns[7]. The only strongly hybridizing fragment (fractions 211–219 in Fig. 1 of ref. 7), designated L7, was cloned[18]. V-L7 is the presumptive germ-line equivalent of V-T2. Also, for V-L7 three independently isolated clones had identical restriction patterns. Subclones of L6 and L7 were constructed (Figs 1, 3) and the pertinent regions sequenced (Figs 2, 4).

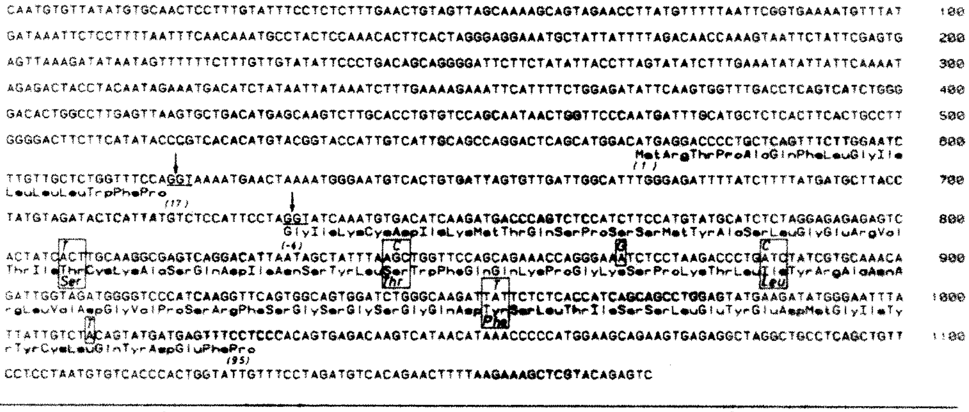
The 5′-flanking region of V-L6 as far as it was sequenced, the leader sequence, the small intron and the first 21 codons of the light-chain gene were found to be identical to the corresponding regions of V-T1 (1–806 in Fig. 2 and Fig. 2 of ref. 17). Further downstream, six differences were found, four of which also coded for different amino acids (Fig. 2). Three of the amino acid differences are located in framework regions and one occurs in the first hypervariable region[19]. Sequence comparisons between V-T1 (ref. 17) and J5 (refs 8,9) show that V-L6 terminates within the Pro codon 95 (position 1,027, Figs 2, 5*a*). In the 3′-flanking region the hepta- and decanucleotide sequences which are complementary to sequences upstream of the J-gene segments[8,9] are found in the expected positions (1,031–1,037; 1,049–1,058).

The 5′-flanking region of V-L7, the leader sequence, the small intron and the first 49 codons of the light-chain gene are identical to the corresponding regions of V-T2 (1–466 in Fig. 4 and 755–1,220 in Fig. 5 of ref. 17). According to restriction patterns the homology extends further upstream[16,18] (at least 6.2 kilobases from the leftward *Sau*3A site in Fig. 3). V-L7 itself differs from V-T2 in 6 base pairs (bp) (Fig. 4), resulting in two different amino acid assignments in the framework region and three in the hypervariable regions. The conserved hepta- and decanucleotide boxes on the 3′ side are also found in V-L7 in the expected positions; therefore, the aberrant joining found in V-T2 (ref. 17) cannot be explained by unusual features of these boxes. The joining point and, with that, the end of the coding region of V-L7, cannot be determined exactly because three different recombination events between CCAACCACA of V-L7 (positions 602–610) and GTGTACACG of J2 would lead to the same recombined sequence CCAACACGT in V-T2 (Fig. 5*b*). The observed reading frameshift in the J-gene segment of V-T2 (ref. 17) can thus be explained by deletion of T, A or C of the J2 sequence.

Although L6/T1 and L7/T2 cannot be assigned to a known subgroup or isotype, an amino acid sequence derived from the V-L6 sequence is identical to that expressed in the tumour RPC-5 which has been determined up to position 38 (S. Rudikoff, personal communication). V-L6 may therefore be the germ-line gene from which the V-gene segment of RPC-5 was formed without mutations within the first 38 codons, whereas on formation of V-T1 two such mutations were introduced in this region.

V-L6 and V-L7 have been called here the presumptive germ-line equivalents of V-T1 and V-T2, respectively. Are they really the germ-line equivalents? We cannot exclude the theoretical possibility that pseudoallelism[20] is the reason for the sequence differences; one would have to assume that the 'true' germ-line equivalents of V-T1 and V-T2 with the respective identical sequences were deleted from the germ line of our BALB/c mice after myeloma T has been induced, and the hybridization experiments with the V-T1 and V-T2 probes therefore show V-L6 and V-L7, respectively, to be the most closely related genes. On the basis of blot hybridization (M.P., unpublished results) it is unlikely (but cannot be excluded) that myeloma T, although it is propagated on BALB/c mice, is derived from another mouse strain which has undergone mutations in the

Fig. 2 Nucleotide sequence of the V-gene region of L6. The sequence extends at the 5′ end across the *Sau*3AI site (see Fig. 1) but numbering was made consistent with Fig. 2 of ref. 17. Differing nucleotides or amino acids in V-T1 are indicated as superscripts or subscripts, respectively. The 3′-flanking region starts at position 1,028.



```
CAATGTGTTATATGTGCAACTCCTTTGTATTTCCTCTCTTTGAACTGTAGTTAGCAAAAGCAGTAGAACCTTATGTTTTTAATTCGGTGAAAATGTTTAT  100
GATAAATTCTCCTTTTAATTTCAACAAATGCCTACTCCAAACACTTCACTAGGGAGGAAATGCTATTATTTTAGACAACCAAAGTAATTCTATTCGAGTG  200
AGTTAAAGATATAATAGTTTTTTCTTTGTTGTATATTCCCTGACAGCAGGGGATTCTTCTATATTACCTTAGTATATCTTTGAAATATATTATTCAAAAT  300
AGAGACTACCTACAATAGAAATGACATCTATAATTATAAATCTTTGAAAAGAAATTCATTTTCTGGAGATATTCAAGTGGTTTGACCTCAGTCATCTGGG  400
GACACTGGCCTTGAGTTAAGTGCTGACATGAGCAAGTCTTGCACCTGTGTCCAGCAATAACTGGTTCCCAATGATTTGCATGCTCTCACTTCACTGCCTT  500
GGGGACTTCTTCATATACCCGTCACACATGTACGGTACCATTGTCATTGCAGCCAGGACTCAGCATGGACATGAGGACCCCTGCTCAGTTTCTTGGAATC  600
                                                                              MetArgThrProAlaGlnPheLeuGlyIle
                                                                             (1)
TTGTTGCTCTGGTTTCCAGGTAAAATGAACTAAAATGGGAATGTCACTGTGATTAGTGTTGATTGGCATTTGGGAGATTTTATCTTTTATGATGCTTACC  700
LeuLeuLeuTrpPhePro
                  (17)
TATGTAGATACTCATTATGTCTCCATTCCTAGGTATCAAATGTGACATCAAGATGACCCAGTCTCCATCTTCCATGTATGCATCTCTAGGAGAGAGAGTC  800
                                GlyIleLysCysAspIleLysMetThrGlnSerProSerSerMetTyrAlaSerLeuGlyGluArgVal
                               (-4)
ACTATCACTTGCAAGGCGAGTCAGGACATTAATAGCTATTTAAGCTGGTTCCAGCAGAAACCAGGGAAATCTCCTAAGACCCTGATCTATCGTGCAAACA  900
ThrIleThrCysLysAlaSerGlnAspIleAsnSerTyrLeuSerTrpPheGlnGlnLysProGlyLysSerProLysThrLeuIleTyrArgAlaAsnA
GATTGGTAGATGGGGTCCCATCAAGGTTCAGTGGCAGTGGATCTGGGCAAGATTATTCTCTCACCATCAGCAGCCTGGAGTATGAAGATATGGGAATTTA  1000
rgLeuValAspGlyValProSerArgPheSerGlySerGlySerGlyGlnAspTyrSerLeuThrIleSerSerLeuGluTyrGluAspMetGlyIleTy
TTATTGTCTACAGTATGATGAGTTTCCTCCCACAGTGAGACAAGTCATAACATAAACCCCCATGGAAGCAGAAGTGAGAGGCTAGGCTGCCTCAGCTGTT  1100
rTyrCysLeuGlnTyrAspGluPhePro
                           (95)
CCTCCTAATGTGTCACCCACTGGTATTGTTTCCTAGATGTCACAGAACTTTTAAGAAAGCTCGTACAGAGTC
```

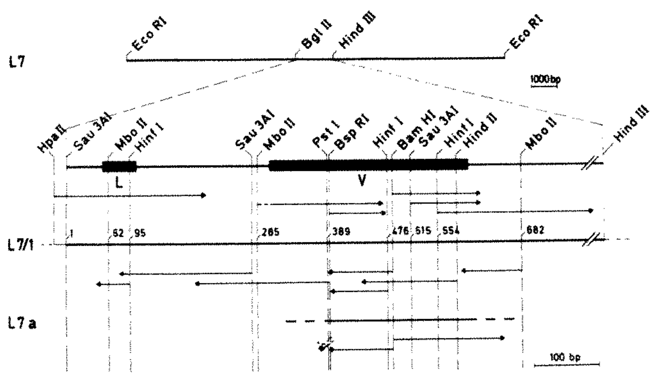


Fig. 3 Partial restriction maps and sequencing strategies of L7 and subclone L7/1. The subclone L7/1 was constructed by ligation of a *Bgl*II/*Hin*dIII digest with a *Bam*HI/*Hin*dIII digest of pBR322 DNA. The restriction map of the second independently isolated λ Charon 4A clone L7a (bottom line) is only shown for the region sequenced. Experimental conditions and presentation of the data are as in Fig. 1.

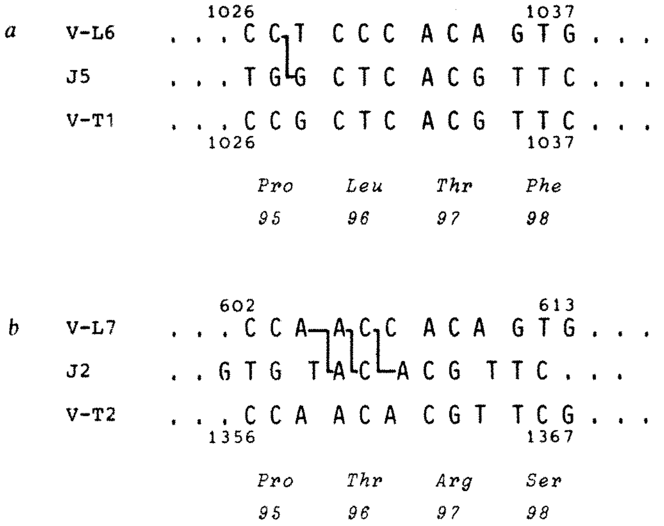

a
1026
1037
V-L6 . . . C C T C C C A C A G T G . . .
J5 . . . T G G C T C A C G T T C . . .
V-T1 . . . C C G C T C A C G T T C . . .
1026
1037
Pro Leu Thr Phe
95 96 97 98
b
602
613
V-L7 . . . C C A A C C A C A G T G . . .
J2 . . G T G T A C A C G T T C . . .
V-T2 . . . C C A A C A C G T T C G . . .
1356
1367
Pro Thr Arg Ser
95 96 97 98

Fig. 5 Joining of V-L6 to J5 at position 1,027 (*a*) and V-L7 to J2 (*b*). In the latter case, joining at positions 604, 605 or 606 leads to the same (non-functional) sequence (see text). Numbering in V-L6 and V-L7 is as in Figs 2 and 4, in V-T1 and V-T2 as in ref. 17; the J2 and J5 sequences are according to refs 8,9.

Fig. 4 Nucleotide sequence of the V-gene region of L7. Nucleotide positions 1–604/606 correlate to nucleotide positions 755–1,358/1,360 of V-T2 in Fig. 5 of ref. 17. Two alternative sites for the L–V gene splicing are indicated: splicing the leader codon 17 with the V-gene codon −1 would produce Gly 17, which is considered an invariant amino acid in light-chain leader sequences (see also ref. 17), while splicing with the codon −4 would produce an Ala 17 but conform better to the splice site consensus sequence and to regularities in leader length. The 3′-flanking region starts at 605/607 (see text). Differing nucleotides or amino acids are indicated as in Fig. 2.

```
GATCTTTCTCAGGGATGTGTCATGGTCCACACAAACTCAGGGAAAGTTTGAAGATGGTATCCACACCTCAGTTCCTTGTATTTTTGCTTTTCTGGATTCC   100
                                                     MetValSerThrProGlnPheLeuValPheLeuLeuPheTrpIlePr
                                                     (1)
 ↓
AGGTATGACTGTTTGGGTGTGGCAAAAAAGTGGAGATGTTATTTAAATACAAAATTTTCTTGCTTTATTTGGAAGCCAATGTCACATGGGAATTGACTTT   200
o(17)
CAGTTTAAAGAAATTGATACAATAAAAGTCATTTATTTTTCTAAGTTGTTTAGAAGTGACTTTCATATTCAGTGTTGTGATCTACTAATGTATCTTCTAT   300
       ↓        ↓
TTTTCCAGCCTCCAGAGGTGACATCTTGCTGACTCAGTCTCCAGCCATCCTGTCTGTGAGTCCAGGAGAAAGAGTCAGTTTCTCCTGCAGGGCCAGTCAG   400
       AlaSerArgGlyAspIleLeuLeuThrGlnSerProAlaIleLeuSerValSerProGlyGluArgValSerPheSerCysArgAlaSerGln
       (-4)     (-1)
                                                                  A             C
AGCATTGGCACAAGCATACACTGGTATCAGCAAAGAACAAATGGTTCTCCAAGGCTTCTCATAAAGTATGCTTCTGAGTCTATCTCTGGGATCCCTTCCA   500
SerIleGlyThrSerIleHisTrpTyrGlnGlnArgThrAsnGlySerProArgLeuLeuIleLysTyrAlaSerGluSerIleSerGlyIleProSerA
                                                                  Asn
                                    C                                    A                  T    G
GGTTTAGTGGCAGTGGATCAGGGACAGATTTTACTCTTAGCATCAACAGTGTGGAGTCTGAAGATATTGCAGATTATTACTGTCAACAAAGTAATAGCTG   600
rgPheSerGlySerGlySerGlyThrAspPheThrLeuSerIleAsnSerValGluSerGluAspIleAlaAspTyrTyrCysGlnGlnSerAsnSerTr
                                   Pro                                 Glu                 TyrArg
GCCAACCACAGTGATGCAGACCATAGCAAAAATCACCAGGAAAGCAGAAGGCTGGGTTGCTTTAGCTATTTCTGGTGGTGTCTCAAACTTCTTCAGAGTG   700
pPro
 (95)
TTTCACAGATGCAGCCATGTGTGATGTTTACAAAAATATGTGGAATGAAGAAGCCCTAAAGTTCTGCTTGACTGTAACATCTTTGCATTCTA
```

germ-line genes corresponding to T1 and T2. However, the simplest interpretation of our blot hybridizations and restriction mapping is the possibility that we are dealing here with the rearranged and non-rearranged counterparts.

Sequence differences have also been explained by residual allele heterogeneity or polymorphism among individual BALB/c mice[10,14]. However, the finding of six sequence differences each within 203 and 131 bp, respectively, and of no differences in the large adjacent regions or in the C_κ gene region[21], argue against polymorphism as the main cause. For the same reason it is unlikely that mutations as they were observed in aged spinner cultures of MOPC-21 (ref. 22) and as they may also occur during tumour propagation contribute to the differences between our germ-line and rearranged sequences.

The relevant regions of V-L6 and V-L7 were sequenced in two independently isolated clones each (Figs 1, 3). No sequence differences were detected. Also, the V-T2 sequence was based on three independent clones[17]. This argues strongly against the supposition that the observed sequence differences arose from cloning artefacts or from fortuitously picking single clones from families of fragments with identical restriction patterns but slightly differing sequences, although such possibilities may never be completely excluded.

The differences between the germ-line and the rearranged V_κ genes could be generated by a somatic mutation mechanism acting preferentially on the coding sequences (perhaps concomitantly with the V–J rearrangement) or by the absence in those sequences of a surveillance process operating elsewhere. It is clear that such a mechanism is not acting on all V_κ genes because Leder's group found the rearranged V_κ gene sequences of MOPC-41 (ref. 12) and MOPC-173B (ref. 13) to be identical to germ-line sequences. The results reported here, on the other hand, add to the known cases of somatic mutations investigated at the DNA[10,14,15] and protein[23] level, a particularly clear case which, because of the clustering of several sequence differences in a small region and the absence of any differences in fairly large adjoining regions, argues in favour of a localized mutation mechanism. Further work must show to what extent such a mutation mechanism contributes to the generation of a functional antibody repertoire.

We thank R. Gerl for assistance, P. S. Neumaier for the computer programs, and particularly S. Rudikoff for communicating his unpublished sequence data. *Bsp*RI, *Hpa*II, *Mbo*II, *Sau*3AI and *Taq*I were from H. Feldmann, W. Gebhard and R. E. Streeck, and *Alu*I and *Bam*HI from Boehringer Mannheim. The work was supported by Bundesministerium für Forschung und Technologie.

Received 5 March; accepted 26 May 1981.

1. Tonegawa, S., Brack, C., Hozumi, N., Matthyssens, G. & Schuller, R. *Immun. Rev.* **36,** 73–94 (1977).
2. Seidman, J. G., Leder, A., Nau, M., Norman, B. & Leder, P. *Science* **202,** 11–17 (1978).
3. Robertson, M. *Nature* **290,** 625–627 (1981).
4. Williamson, A. R. *A. Rev. Biochem.* **45,** 467–500 (1976).
5. Potter, M. *Adv. Immun.* **25,** 141–211 (1977).
6. Rabbitts, T. H., Matthyssens, G. & Hamlyn, P. H. *Nature* **284,** 238–243 (1980).
7. Steinmetz, M., Höchtl, J., Schnell, H., Gebhard, W. & Zachau, H. G. *Nucleic Acids Res.* **8,** 1721–1729 (1980).
8. Sakano, H., Hüppi, K., Heinrich, G. & Tonegawa, S. *Nature* **280,** 288–294 (1979).
9. Max, E. E., Seidman, J. G. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **76,** 3450–3454 (1979).
10. Early, P., Huang, H., Davis, M., Calame, K. & Hood, L. *Cell* **19,** 981–992 (1980).
11. Rudikoff, S., Rao, D. N., Glaudemans, C. P. J. & Potter, M. *Proc. natn. Acad. Sci. U.S.A.* **77,** 4270–4274 (1980).
12. Seidman, J. G., Max, E. E. & Leder, P. *Nature* **280,** 370–375 (1979).
13. Max, E. E., Seidman, J. G., Miller, H. & Leder, P. *Cell* **21,** 793–799 (1980).
14. Bernard, O., Hozumi, N. & Tonegawa, S. *Cell* **15,** 1133–1144 (1978).
15. Sakano, H., Maki, R., Kurosawa, Y., Roeder, W. & Tonegawa, S. *Nature* **286,** 676–683 (1980).
16. Steinmetz, M. & Zachau, H. G. *Nucleic Acids Res.* **8,** 1693–1707 (1980).
17. Altenburger, W., Steinmetz, M. & Zachau, H. G. *Nature* **287,** 603–607 (1980).
18. Höchtl, J. thesis, Univ. München (1980).
19. Kabat, E. A., Wu, T. T. & Bilofsky, H. *NIH Publ.* 80-2008 (1979).
20. Hilschmann, N. *et al. Naturwissenschaften* **65,** 616–639 (1978).
21. Altenburger, W., Neumaier, P. S., Steinmetz, M. & Zachau, H. G. *Nucleic Acids Res.* **9,** 971–981 (1981).
22. Adetugbo, K., Milstein, C. & Secher, D. S. *Nature* **265,** 299–304 (1977).
23. Gearhart, P. J., Johnson, N. D., Douglas, R. & Hood, L. *Nature* **291,** 29–34 (1981).
24. Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65,** 499–560 (1980).

Bleomycin-resistant DNA synthesis in ataxia telangiectasia cells

Philippe Cramer* & Robert B. Painter†

Laboratory of Radiobiology, University of California, San Francisco, California 94143, USA

Ataxia telangiectasia (AT) is a human autosomal recessive disease that exhibits (in addition to the symptoms implied by its name) immune deficiencies, greatly increased risk of lymphoreticular cancer and unusual sensitivity to ionizing radiation (see ref. 1 for review). Although cells from AT patients are abnormally sensitive to killing by X-ray irradiation, DNA synthesis in AT cells is resistant to such treatment[2,3]. This resistance is primarily due to less inhibition of replicon initiation in AT cells than in normal cells, but there is also less inhibition of chain elongation in AT cells[3]. We now report that DNA synthesis in AT cells is also resistant to DNA damage induced by bleomycin but not to damage induced by *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine (MNNG) or actinomycin D. Of these agents, only X rays and bleomycin directly cause DNA strand breaks, indicating that AT cells are defective in the system that in normal cells prolongs the time during which strand breaks can be repaired, replicon initiation inhibited and the chromatin structure returned to its normal state.

Normal human fibroblasts (HS27) and AT3BI cells were exposed for 30 min to various concentrations of bleomycin and their rates of DNA synthesis compared 30 min later (Fig. 1). The dose response for inhibition of DNA synthesis in normal cells showed a steep decline at low doses and a shallow decline at high doses, as previously reported[4]. It is generally assumed that the steep component is primarily due to inhibition of replicon initiation[5-7] and the shallow one to blocked chain elongation in individual replicons[8]. The dose response for AT3BI cells, which are classified as excision deficient[1], showed no decrease in rate of synthesis at concentrations up to 75 μg ml^{-1}; at higher concentrations DNA synthesis was inhibited, but to a much lesser extent than in normal cells. In AT5BI cells, which are classified as excision proficient[4], almost identical results were obtained (data not shown). These responses are similar to those observed with X-ray irradiation[1,2].

Sedimentation analysis of DNA on alkaline sucrose gradients confirmed the interpretation that bleomycin-induced DNA damage that inhibited replicon initiation in normal cells failed to do so in AT cells (Fig. 2). The major difference between the profiles of newly synthesized DNA from bleomycin-treated AT3BI cells and normal human fibroblasts is that, for the latter at 30 min after bleomycin treatment, there was a much larger reduction of radioactivity in low-molecular-weight regions (fractions 4–10) compared with untreated cells. This radioactivity reflects the part of DNA synthesis due to replicon initiation[7-9]. (That inhibition of replicon initiation is the principle basis for the inhibition of DNA synthesis induced by low doses of X rays (< 1,000 rad) has been confirmed by autoradiography[10] and *p*H step alkaline elution[9].) Chain elongation was also more resistant to bleomycin in AT3BI than in normal cells. In treated and untreated AT cells the high-molecular-weight regions of the profiles (fractions 13–24), which reflect chain elongation and joining of newly formed DNA[7-9], were almost superimposable, whereas the same regions of the profiles of DNA from treated normal cells show a reduction of radioactivity. These results cannot be explained on the basis of reduced entry of bleomycin into AT cells, because Taylor *et al.*[11] have shown that equal concentrations of bleomycin in the medium induce more chromosomal aberrations in AT cells than in normal cells. Moreover, preliminary data (not shown) indicate that equal numbers of single-strand DNA breaks are induced by bleomycin in AT and normal cells. Thus, AT cells are resistant to the bleomycin-induced damage that inhibits replicon initiation and blocks chain elongation in normal human cells.

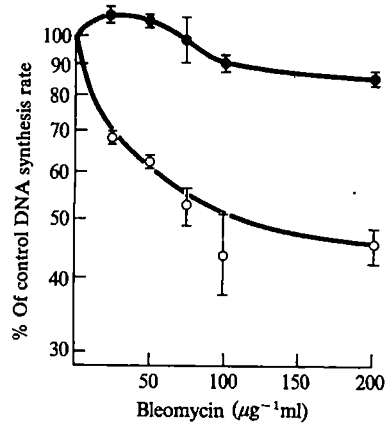


Fig. 1 Effect of bleomycin on rate of DNA synthesis in normal diploid HS27 (○) and AT3BI (●) cells. Cells grown in Eagle's minimal essential medium supplemented with 15% fetal bovine serum were inoculated into 35-mm Petri dishes at ~ 5×10^4 cells per dish. After 1 day of growth the medium on all cultures was replaced with medium containing 0.01 μCi ml^{-1} of ^{14}C-thymidine (50 mCi mmol^{-1}) and incubation was continued for 24 h. The medium was removed and two dishes containing AT cells and two containing HS27 cells were incubated for 30 min in medium containing various concentrations of bleomycin, after which this medium was replaced with nonradioactive medium and the cells returned to the incubator for 30 min. The medium was then removed and all cultures were incubated with medium containing 20 μCi ml^{-1} of ^{3}H-thymidine (50 Ci mmol^{-1}) for 10 min; the cells were washed with ice-cold SSC (0.15 M sodium chloride, 0.015 M sodium citrate) and scraped into SSC. Ice-cold 4% perchloric acid was added and the cell suspension filtered through GF/C (Whatman) filters, which were rinsed sequentially with 4% perchloric acid, 70% alcohol and 100% alcohol, dried, and assayed for radioactivity in a liquid scintillation spectrometer. The resulting ^{3}H/^{14}C ratios were measures of the specific activity of DNA and therefore of the rate of DNA synthesis. Error bars indicate the range of duplicate samples.

As has been suggested for X-ray irradiation[2,3], the reason for the observed differences in DNA replication after bleomycin treatment may be that AT cells lack a factor or process that in normal cells intervenes to delay replication of damaged regions of DNA, or that AT cells have an altered chromatin structure that does not respond to bleomycin-induced damage. Previous studies have suggested that, in the chromatin of normal cells, DNA strand breaks induce a conformational change that inhibits initiation of DNA synthesis throughout a replicon cluster[12]. If this is true, AT chromatin must not experience this conformational change after damage induced by bleomycin.

To determine whether the reported sensitivities of AT cells to killing by MNNG and actinomycin D[13,14] are correlated with those of bleomycin, we measured the inhibition of DNA synthesis induced by these two agents (Fig. 3, Table 1). In neither case was there a significant difference between the two cell types in the effects of these agents on DNA synthesis.

Thus we suggest that high yields of DNA strand breaks are directly produced by X rays and bleomycin[15] and are responsible for the inhibition of initiation of replicon clusters in normal cells. Actinomycin D and MNNG do not directly produce many strand breaks and therefore do not affect DNA synthesis at the level of the cluster, but instead induce lesions (methylated bases and intercalated molecules) that inhibit DNA replication primarily at the level of the individual replicon in both normal and AT cells. However, some of the 'X-ray-like' damage is induced by these agents. In AT cells, the defective damage-recognition system fails to respond to this damage and allows DNA

---

* Present address: Institut de Recherche d'Hématologie et d'Oncologie, Département de Pédiatrie, Clinique des Maladies du Sang, Hôpital Saint-Louis, 2, Place du Docteur Fournier, 75010 Paris, France.

† To whom correspondence should be addressed.

0028-0836/81/260671—02$01.00

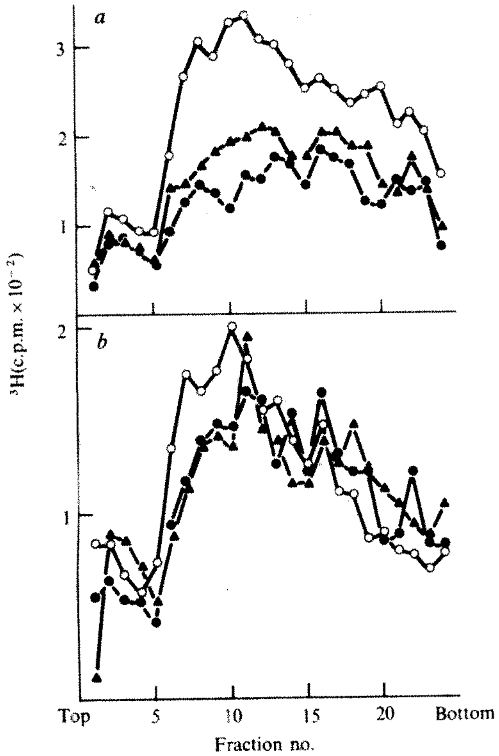


Fig. 2 Alkaline sucrose gradient profiles of newly formed DNA from normal diploid HS27 (*a*) and AT3BI (*b*) cells 30 min after treatment with bleomycin. ○, Untreated cells; ▲, $25\ \mu g\ ml^{-1}$ bleomycin; ●, $50\ \mu g\ ml^{-1}$ bleomycin. After the 30-min incubation in drug-containing medium, the cells were incubated in drug-free medium for 30 min and then in medium containing ^{3}H-thymidine for 10 min. The cells were immediately cooled to ice temperatures, suspended in ice-cold buffer, irradiated with 1,000 rad to aid untangling of strands and lysed directly on top of 5–20% alkaline sucrose gradients. Centrifugation for 3 h at 27,000 r.p.m. in a Beckman SW-27 rotor was followed by collection of fractions, counting of radioactivity and analysis as described previously[7].

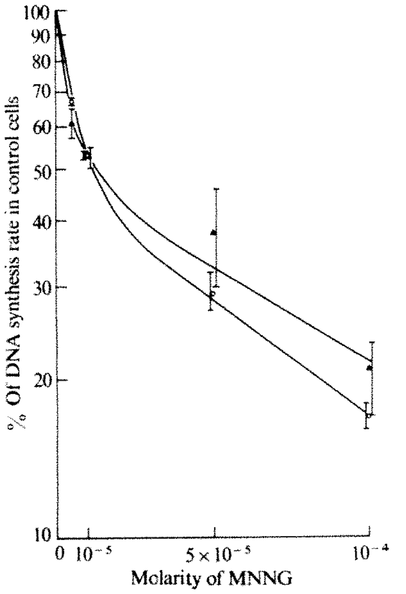


Fig. 3 Effect of MNNG on rate of DNA synthesis in normal diploid HS27 (○) and AT3BI (▲). The procedure is the same as that used with bleomycin (see Fig. 1 legend). Error bars indicate the range of duplicate samples.

Table 1 Inhibition of DNA synthesis by actinomycin D in AT and normal cells

Concentration of actinomycin D (M)	DNA synthesis (% of control) Normal cells	AT cells
5.0×10^{-7}	76 ± 1.0	76 ± 2.0
7.5×10^{-7}	60 ± 5.0	72 ± 2.0
1.0×10^{-6}	47 ± 2.0	54 ± 8.0
2.0×10^{-6}	40 ± 1.0	21 ± 5.0
5.0×10^{-6}	11 ± 0.5	9 ± 1.0

The procedure was as used with bleomycin (see Fig. 1).

replication to occur on damaged templates. This defective system also allows cells to progress from G_2 into mitosis before pre-aberrational damage is completely repaired, resulting in the reduced mitotic delay[16] and increased frequencies of chromosomal aberrations[17].

This hypothesis would not seem to explain the enhanced frequencies of chromosomal aberrations observed after irradiation of AT lymphocytes in G_0 (ref. 17), that is, before stimulation by a mitogen. However, Saha and Tolmach[18] showed that HeLa DNA synthesis was inhibited on entry into S phase even when the cells had been irradiated several hours earlier while in G_1. If this phenomen is due to a delay during which pre-aberrational lesions are repaired, the defective damage-recognition system in AT cells may cause them to progress precociously through S phase before repair is complete and thus allow the eventual expression of the lesion in mitosis.

This work was supported by the US Department of Energy and by a grant from the Fondation de l'Industrie Pharmaceutique pour la Recherche, Paris.

Received 31 October 1980; accepted 10 April 1981.

1. Paterson, M. C. & Smith, P. J. *A. Rev. Genet.* **13,** 291–318 (1979).
2. Houldsworth, J. & Lavin, M. F. *Nucleic Acids Res.* **8,** 3709–3720 (1980).
3. Painter, R. B. & Young, B. R. *Proc. natn. Acad. Sci. U.S.A.* **77,** 7315–7317 (1980).
4. Watanabe, M., Takabe, Y., Katsumata, T., Terasima, T. & Umezawa, H. *J. Antibiot., Tokyo* **26,** 417–423 (1973).
5. Makino, F. & Okada, S. *Radiat. Res.* **62,** 37–51 (1975).
6. Walters, R. A. & Hildebrand, C. E. *Biochem. biophys. Res. Commun.* **65,** 265–271 (1975).
7. Painter, R. B. & Young, B. R. *Biochim. biophys. Acta* **418,** 146–153 (1976).
8. Watanabe, I. *Radiat. Res.* **58,** 541–556 (1974).
9. Erickson, L. C., Ross, W. E. & Kohn, K. W. *Chromosoma* **74,** 125–139 (1979).
10. Dahle, D., Griffiths, T. D. & Carpenter, J. G. *Radiat. Res.* **78,** 542–549 (1979).
11. Taylor, A. M. R., Rosney, C. M. & Campbell, J. B. *Cancer Res.* **39,** 1046–1050 (1979).
12. Povrik, L. F. & Painter, R. B. *Biochim. biophys. Acta* **432,** 267–272 (1976).
13. Scudiero, D. A. *Cancer Res.* **40,** 984–990 (1980).
14. Hoar, D. I. & Sargent, P. *Nature* **261,** 590–592 (1976).
15. Suzuki, H., Nagai, K., Yamaki, H., Tanaka, N. & Umezawa, H. *J. Antibiot., Tokyo* **22,** 446–448 (1969).
16. Zambetti-Bosseler, F. & Scott, D. *Int. J. Radiat. Biol.* (in the press).
17. Taylor, A. M. R., Metcalfe, J. A., Oxford, J. M. & Harnden, D. G. *Nature* **260,** 441–443 (1976).
18. Saha, B. K. & Tolmach, L. J. *Radiat. Res.* **66,** 76–89 (1976).

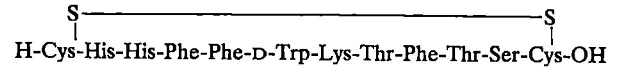
Absence of luteinizing hormone-releasing and anti-fertility properties in a glucagon-selective somatostatin analogue

Frederick J. Bex*, Alan Corbin*, Dimitrios Sarantakis† & Eric Lien‡

Sections of Endocrinology*, Medicinal Chemistry†
and Biochemistry‡, Wyeth Laboratories, Inc., Box 8299,
Philadelphia, Pennsylvania 19101, USA

A dodecapeptide analogue of somatostatin (SRIF), [des-Ala^1, Gly^2]His^{4,5}-D-Trp^8-SRIF (Wy-41,747), combining selective inhibition of growth hormone and glucagon release with prolonged activity, prompted its consideration as an adjunct in the treatment of diabetes mellitus[1,2]. Chronic treatment of mature rats with Wy-41,747 produced reproductive effects resembling those described for agonistic analogues of the luteinizing hormone releasing hormone (LHRH). The integral properties of the LHRH agonists such as LH release, induction of ovulation, and the ability to inhibit pregnancy both before and after implantation were noted[3]. We report here that the above pharmacological findings were observed only with a single batch (Wy-41,747E) which was accidentally contaminated with 0.023–0.06% by weight of an LHRH analogue (Wy-40,972).[4]

A rigorous examination of Wy-41,747E and of a second batch of Wy-41,747, designated H, was initiated after a personal communication to A.C. by J. Rivier and W. Vale concerning the absence of *in vitro* LH release with a sample of Wy-41,747 prepared and tested at the Salk Institute. The E and H batches were bioequivalent with respect to insulin and glucagon suppression in a standard procedure[5] using male rats pretreated with Nembutal and arginine. The batches, however, were clearly differentiated by their ovulation induction activity, itself a measure of LH-releasing activity, in the Nembutal-treated pro-estrous rat[3]. The E batch induced ovulation at subcutaneous doses of 25–100 μg per rat, and such effects were dose-related, whereas Wy-41,747H had no activity at a dose of 200 μg per rat. Wy-40,972 was suspected as the most likely contaminant due to an earlier involvement in a large-scale synthesis of this potent LHRH agonist. The structures of Wy-41,747 and Wy-40,972 are:



$$\overbrace{\text{H-Cys-His-His-Phe-Phe-D-Trp-Lys-Thr-Phe-Thr-Ser-Cys-OH}}^{\text{S———————————S}}$$

Wy-41,747

pGlu-His-Trp-Ser-Tyr-D-Trp-N-CH_3Leu-Arg-Pro-NH-CH_2-CH_3

Wy-40,972

Analytical HPLC indicated that the k' (retention volume) of a 0.023% impurity in Wy-41,747E was identical to the k' of Wy-40,972. The impurity was not detected in Wy-41,747H. Comparative binding studies in a radioimmunoassay specific for Wy-40,972 confirmed the identity of the contaminant. Wy-41,747H does not bind to the Wy-40,972 antibody, but the Wy-41,747H samples spiked with Wy-40,972 do bind to a degree related to the amounts of Wy-40,972 present and do induce ovulation in the expected manner. Binding for Wy-41,747E was evident and correlated with a 0.06% weight content of Wy-40,972.

The incident resulting in contamination has not been identified. This is the second reported case of accidental contamination with minute amounts of peptide LH-releasing agents[6]. Other investigators in this field should be cautioned to take extreme care with equipment and media previously exposed to these powerful hormones.

We conclude that the reproductive effects seen in toxicological and endocrine studies with Wy-41,747E are attributable to contamination with Wy-40,972. Wy-41,747 remains of interest as a potentially useful somatostatin analogue.

We thank Drs Jean Rivier and W. Vale of the Salk Institute, California, for drawing our attention to a possible contaminant.

Received 12 March; accepted 30 March 1981.

1. Lien, E. L., Greenwood, J. & Sarantakis, D. *Diabetes* **28,** 491–495 (1979).
2. Sarantakis, D., Teichman, J., Fenichel, R. & Lien, E. L. *FEBS Lett.* **92,** 153–155 (1978).
3. Bex, F. J., Corbin, A., Sarantakis, D. & Lien, E. L. *Nature* **284,** 342–343 (1980).
4. Foell, T. J. & Rees, R. W. U.S. Patent 4,089,946 (1978).
5. Grant, N. *et al. Life Sci.* **19,** 629–632 (1976).
6. Rivier, J. *et al. J. med. Chem.* **16,** 545–549 (1973).

Multiple sequences related to classical histocompatibility antigens in the mouse genome

B. Cami, F. Brégégère, J. P. Abastado & P. Kourilsky

Unité de Biologie Moléculaire du Gène, ER CNRS 201 and SCN INSERM 20, Institut Pasteur, 28 rue du Dr Roux, 75724 Paris Cédex 15, France

A series of genes on the 17th chromosome of the mouse, called the major histocompatibility complex, has critical roles in the immune response. The classical transplantation antigens (or H–2 antigens) are encoded by genes located in several loci of this complex—*H-2K, H-2D* and *H-2L* (refs 1, 2; Fig. 1). The antigens are cell-surface glycoproteins consisting of a single polypeptide chain of about 350 amino acids, non-covalently associated to a shorter polypeptide chain, β_2-microglobulin, which is not encoded by the 17th chromosome. Within a given *H-2* haplotype, H-2K, -D or -*L* antigens differ by discrete antigenic sites, probably due to changes in their amino acid sequences[3]. As many as 50 alleles have been found for each *K* and *D* locus in the laboratory mouse; thus *H-2* loci are very polymorphic[2]. We report here a preliminary study of the structure of *H-2* genes, in which the number of *H-2*-related sequences in the mouse genome has been estimated from the number of DNA fragments able to hybridize specifically with a cDNA carrying an *H-2* sequence.

We have previously described the characterization of recombinant plasmids carrying $H\text{-}2^d$-specific cDNA inserts[4,5]. One of them, pH-2^d-3, was used as a source of hybridization probes to estimate the number of *H-2*, or *H-2*-related genes present in the mouse genome. A restriction map of the insert carried by pH-2^d-3 is shown in Fig. 2, together with the corresponding structural regions of the *H-2* heavy chain. pH-2^d-3 contains sequences coding for the third domain, the membrane-spanning region, the cytoplasmic COOH terminus and the 3′-untranslated region down to the poly(A) (Fig. 2).

In a first series of experiments, we used the entire pH-2^d-3 cDNA as a probe for investigating mouse genome. Liver genomic DNA was prepared from BALB/c ($H\text{-}2^d$) animals, digested by *Eco*RI and used in Southern's blotting technique[6]. Smears were observed, which were not caused by poly(A) or by G–C tails, indicating the detection by the probe of sequences highly repeated in the genome. A mouse genomic library was also screened with the same probe. This library was prepared from 15–20-kilobase BALB/c mouse DNA, produced by partial digestion with *Sau*3A and cloned in λ Charon 28 (F. Blattner, personal communication). The rate of positive response was 10–20%.

Screening experiments were then repeated using, in parallel, *Pst*I restriction fragments A, B and C as probes. The 160-base pair fragment C was the only one to hybridize with 30 genomic clones out of 200, whereas the two other probes revealed no positive responders. We conclude, therefore, that fragment C contains a sequence which is repeated about 40,000-fold in the haploid genome of the mouse. Fragment C corresponds to a part of the 3′-untranslated region of the messenger and has been shown to be highly conserved in most of the cDNAs so far sequenced specifying different H–2 alloantigens (in preparation). It thus appears to be an intriguing and original feature of the *H-2* genes that they contain a highly repeated sequence in the 525-base pair, 3′-untranslated region of their mRNA.

To probe specifically for H–2 coding sequences, we prepared a *Bgl*I–*Hin*fI restriction fragment (probe D, Fig. 2). It codes for amino acids 195–283 of H–2 antigen, and corresponds nearly exactly to the third domain of this molecule[5]. Genomic DNAs of various origins were digested by the appropriate restriction enzymes and analysed using Southern's blotting technique[6]. Results are given in Fig. 3.

On hybridization with fragment D, liver BALB/c DNA ($H\text{-}2^d$), digested by *Eco*RI or *Bam*HI, results in about 20 and 15 bands respectively (lanes *a* and *b*). These bands display, reproducibly, different intensities and are distributed between 1.2 and 26 kilobases. This is not peculiar to either liver DNA or BALB/c mice, because liver DNA from C57BL/6 mice ($H\text{-}2^b$), digested by *Eco*RI or *Bam*HI, produced a similar number of bands (lanes *c* and *d*), as did C57BL/6 kidney DNA digested by *Eco*RI (lane *i*). The liver and kidney DNAs prepared from C57BL/6 mice yielded identical patterns (lanes *h* and *i*); in contrast, liver DNAs prepared from mice differing in their *H-2* haplotypes, resulted in different patterns (lanes *a*–*d*). Internal controls (Fig. 3, lanes *e* and *f*) rule out the possibility that these multiple bands are due to partial digestion of mouse DNA. No *Eco*RI or *Bam*HI sites have been found in the cDNA sequences coding for the third domain of three different H–2 alloantigens (ref. 5 and unpublished data). In addition, the 290-base pair probe A corresponds to the third domain of the H–2 molecule (Fig. 2). As compared with other genes[7,8], introns would be

0028-0836/81/260673—03$01.00

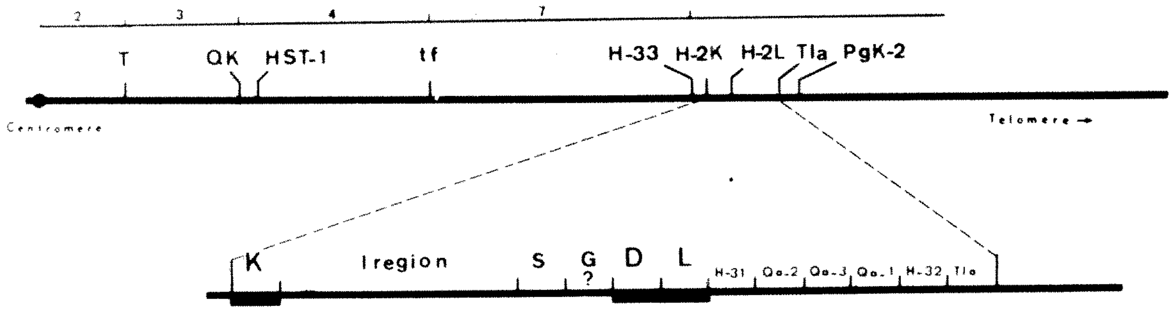


Fig. 1 Top: genetic map of the left-hand part of the 17th chromosome of the mouse, according to ref. 1. The genetic markers indicated are, from left to right: brachyury (T), quaking (QK), hybrid sterility-1 (HST-1), tufted (tf), histocompatibility 33 (H–33), histocompatibility-2K (H–2K), histocompatibility-2L (H–2L), thymusleukaemia antigen (Tla) and phosphoglycerate kinase-2 (PgK-2). Bottom: detailed genetic map of the major histocompatibility complex. Regions K, D and L, which are known to code for classical histocompatibility antigens, are indicated with thick lines. Region I contains the Ir and Ia genes, S codes for protein components of the complement, G has been reported to code for a specific H–2 antigen but its existence is controversial[1]. H–31, Qa–1, –2 and –3, H–32 and Tla are minor histocompatibility antigens, sometimes referred to as 'H–2 like'. The order of the loci coding for them is not known.

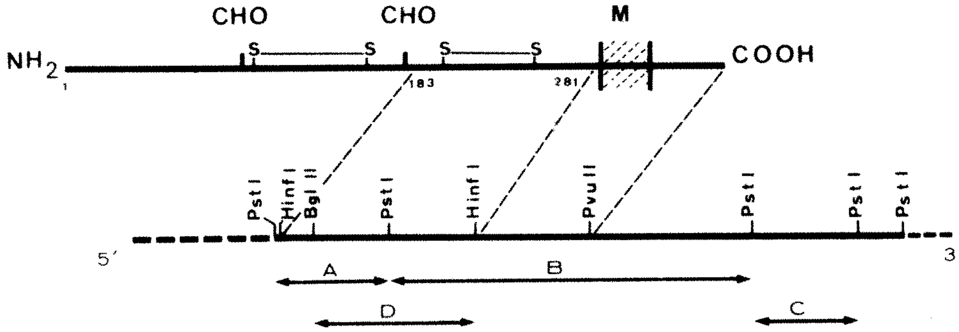


Fig. 2 The restriction map of the cDNA insert of the recombinant plasmid pH-2^d-3 (ref. 5) is presented in correspondence with the linear structure of an H–2 antigen molecule[16]. M indicates the membrane-spanning region with the external part of the molecule on the NH$_2$-terminal side, and the cytoplasmic fragment on the COOH-terminal side. A, B, C and D represent restriction fragments used as radiolabelled probes in the experiments described in the text.

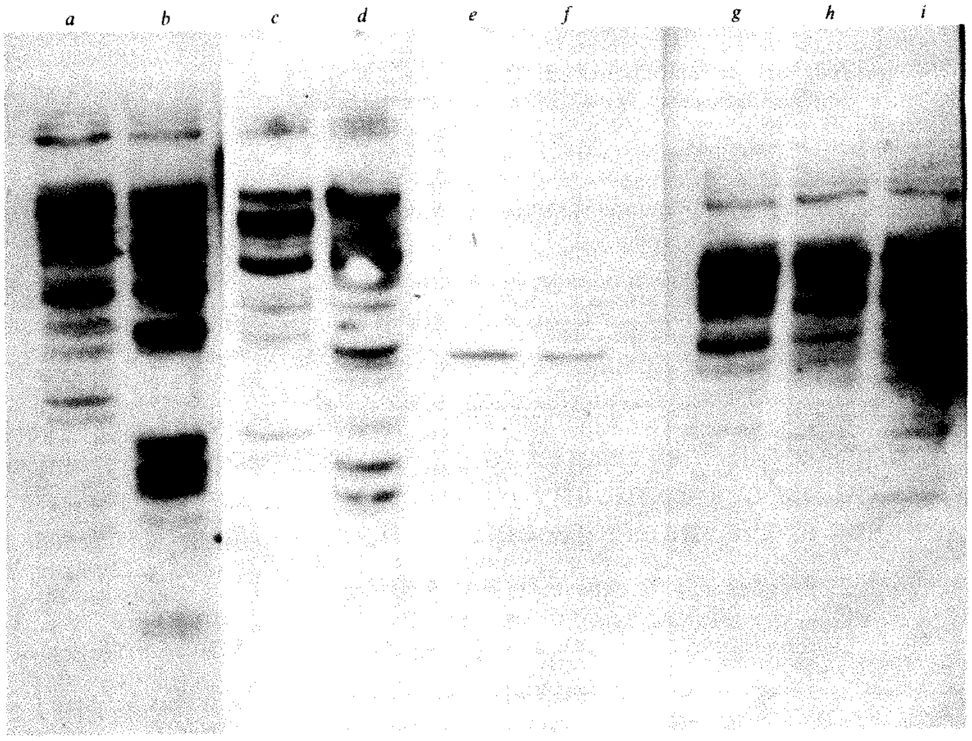


Fig. 3 The high-molecular-weight mouse genomic DNA was extracted and purified as described elsewhere[17]. DNA samples (20 μg) were digested by the appropriate restriction enzymes and subjected to electrophoresis in a 0.7% agarose gel. The DNA was then transferred to nitrocellulose sheets as described by Southern[6]. The sheets were prehybridized for 4 h at 68 °C in 10× Denhardt (1 × Denhardt = 0.02% Ficoll, 0.02% polyvinylpyrrolidone, 0.02% bovine serum albumin), 6 × SSC (1 × SSC = 0.15 M NaCl, 0.015 M Na citrate, *p*H 7). Restriction fragment D from pH-2^d-3 (Fig. 2) was polymerized using T4 DNA ligase (Boehringer) and then ^{32}P-radiolabelled by nick-translation[18] to a specfic activity of 2×10^8 c.p.m. per μg. Each nitrocellulose sheet was hybridized with 100 ng of this probe for 16 h at 68 °C in 1 × Denhardt, 2 × SSC, 25 mM KH_2PO_4, *p*H 7.2 (NaOH), 2 mM EDTA, 0.5% SDS, 50 μg ml^{-1} sonicated herring sperm DNA. The nitrocellulose sheets were then washed four times for 45 min at 68 °C in 1 × Denhardt, 2 × SSC, 0.5% SDS, and twice for 45 min at 68 °C in 0.2 × SSC, 0.1% SDS. They were dried for 10 min at 80 °C and autoradiographed on flash-activated Kodak Royal-XO-Mat films for 3–10 days at −70 °C with intensifying screens. Lanes *a* and *b*, liver DNA from BALB/c mice (H–2^d haplotype) digested by *Eco*RI and *Bam*HI, respectively. Lanes *c* and *d*, liver DNA from C57BL/6 mice (*H–2^b* haplotype) digested by *Eco*RI and *Bam*HI respectively. Lanes *e* and *f*, same as lanes 1 and 2, but 0.5 ng of pH-2^d-3 DNA has been added to mouse DNA before digestion, to provide total digestion controls. Lanes *g*, *h* and *i*, liver DNA from BALB/c, liver and kidney DNAs from C57BL/6 mice, respectively, were digested with *Eco*RI.

expected to occur between, rather than inside, the domains. The presence of *Eco*RI or *Bam*HI sites is then unlikely to occur within the genomic sequences coding for the third domain of H-2 antigens: as a consequence, most of these genomic sequences are expected to give a single hybridization band. Hybridizations were carried out in stringent conditions (Fig. 3 legend): no cross-hybridization of the probe is thus expected to take place with sequences coding for immunoglobulin constant domains or for β_2-microglobulin, although these display some amino acid sequence homologies with the third domain of H-2 antigens[9,10]. Similarly, no cross-hybridization should be detected with sequences coding for the first and second domains of H-2 antigens, which share no apparent homology with the third domain[11,12].

The same mouse genomic library as above was screened, using as a probe fragment B, which partially extends into the third domain (Fig. 2). As cuts in mouse DNA have been generated at random and the size of the inserts is much larger than that of the probe, the frequency of positive clones should provide a measure of the number of *H-2*-related sequences in the mouse genome: about 750,000 recombinants were screened and 65 positive ones isolated. The library was not composed of independent clones and there could have been selective amplification of some of them. However, restriction mapping of the DNA of isolated positive clones showed that more than half of them, at least, were different from one another (data not shown). Because the mouse genome should be represented in about 250,000 recombinants, the results indicate the existence of at least 10–12 sequences reacting with probe B per haploid genome. This figure is most probably an underestimate because many positive recombinants usually escape or are lost during isolation.

Thus, the present data suggest strongly that 15–20 restriction fragments of the mouse genome possess sequences sharing extensive homology with the third domain of H-2 antigens, three to five times more than expected from the genetics of H-2 antigens. Of course, some of these sequences could correspond to pseudogenes, as described in the globin gene family[13,14]. Also, 15–20 *H-2*-related sequences can exist in the mouse genome. A possible explanation for the existence of bands of stronger intensity (Fig. 3) could be the occurrence of several *H-2* (third domain) genomic sequences in the corresponding DNA fragments. On the other hand, note that several cell-surface antigens, such as Tla, Qa1, -2 and other minor transplantation antigens, are composed of a heavy chain of molecular weight 44,000 associated with β_2-microglobulin, a structure like that of H-2 antigens[15]. The sequence homologies reported between the third domain of H-2 molecules, β_2-microglobulin and the constant domains of immunoglobulins suggest that structures similar to the third domain may be conserved in a variety of genes. In particular, this might be expected for genes coding for proteins interacting with β_2-microglobulin itself.

In this respect, the use of a hybridization probe specific for the third domain may have revealed the existence of a multigene family containing, in addition to genuine *H-2K*, *-D* and *-L* genes, perhaps, Tl, Qa1 and/or unpreviously characterized genes encoding H-2-like antigens. Analysis of the isolated genes should answer this question.

We thank Dr F. Blattner for the gift of the mouse genomic DNA library, Dr F. Rougeon for mouse kidney DNA, Miss O. Le Bail for technical assistance, and Dr G. Gachelin for critical reading of this manuscript. The work reported was done in accordance with the French Guidelines for Recombinant Research. It was supported by grants from the CNRS (ER 201, ATP 4246), INSERM (SCN 20, CRAT 72-79-104) and DGRST (79-7-0500, 78-7-2931).

Received 11 March; accepted 8 April 1981.

1. Klein, J. *Science* **203,** 516–521 (1979).
2. Klein, J. *Biology of the Mouse Histocompatibility-2 Complex* (Springer, New York, 1979).
3. Nairn, R., Yamaga, K. & Nathenson, S. G. *A. Rev. Genet.* **14,** 241–277 (1980).
4. Kvist, S. *et al. Proc. natn. Acad. Sci. U.S.A.* **78** (in the press).
5. Brégégère, F. *et al. Nature* (in the press).
6. Southern, E. M. *J. molec. Biol.* **98,** 503–517 (1975).
7. Sakano, H. *et al. Nature* **277,** 627–633 (1979).
8. Eaton, W. A. *Nature* **284,** 183–185 (1980).
9. Tragardh, L., Rask, L., Wiman, K., Fohlman, J. & Peterson, P. A. *Proc. natn. Acad. Sci. U.S.A.* **76,** 5839–5842 (1979).
10. Orr, H. T., Lopez de Castro, J. A., Lancet, D. & Strominger, J. L. *Nature* **282,** 266–270 (1979).
11. Orr, H. T., Lopez de Castro, J. A., Lancet, D. & Strominger, J. L. *Biochemistry* **18,** 5711–5720 (1979).
12. Tragardh, L., Rask, L., Wiman, K., Fohlman, J. & Peterson, P. A. *Proc. natn. Acad. Sci. U.S.A.* **77,** 1129–1133 (1980).
13. Proudfoot, N. J. & Maniatis, T. *Cell* **21,** 537–544 (1980).
14. Lacy, E. & Maniatis, T. *Cell* **21,** 545–553 (1980).
15. Vitetta, E. S. & Capra, J. D. *Adv. Immun.* **26,** 147–193 (1968).
16. Kimball, E. S., Maloy, E. S., Martinko, J. M., Nathenson, S. G. & Coligan, J. E. *Molec. Immun.* **17,** 1283–1291 (1980).
17. Gross-Bellard, M., Oudet, P. & Chambon, P. *Eur. J. Biochem.* **36,** 32–38 (1973).
18. Maniatis, T., Jeffrey, A. & Kleid, D. *Proc. natn. Acad. Sci. U.S.A.* **72,** 1184–1188 (1975).

Homologous tyrosine phosphorylation sites in transformation-specific gene products of distinct avian sarcoma viruses

James C. Neil*, Jacques Ghysdael*, Peter K. Vogt*
& John E. Smart†

* Department of Microbiology, University of Southern California
School of Medicine, 2011 Zonal Avenue, Los Angeles,
California 90033, USA
† Cold Spring Harbor Laboratory, Cold Spring Harbor,
New York 11724, USA

Three classes of avian sarcoma viruses have been defined recently (Table 1)[1–5]. Transformation-specific proteins of all three viral classes have associated tyrosine-specific protein kinase activity[4–12]. The protein kinase activity of Rous sarcoma viruses (RSV), representing class I, seems to be a property of the transformation-specific protein, $pp60^{src}$, itself[13] and seems to be essential for cellular transformation, as the activity is labile or absent in transformation mutants[6,7,14,15]. This connection between transformation-specific protein, enzyme activity and oncogenic cellular changes has yet to be established for the viruses of classes II and III, which code for polyproteins in which *gag*-derived and transformation-specific sequences are fused. We report here an analysis of the phosphorylation sites on the transformation-specific proteins encoded by RSV and the class III viruses which shows that, although there is no appreciable sequence relationship between large portions of these molecules, a tryptic peptide containing the phosphotyrosine seems to be identical in classes I and III.

In this study we made use of the fact that class II and III avian sarcoma virus transformation-specific proteins are labelled in *in vitro* phosphotransfer reactions on immunoprecipitates[3–5,9–12]. Until conclusive experiments using purified viral proteins have been done, the possibility remains that these reactions are catalysed by cell-coded proteins associated with the immunoprecipitated material. Note that there are conflicting reports on

Table 1 Avian sarcoma viruses

Class	Virus	Transformation-specific protein
I	RSV	$pp60^{src}$
	Avian sarcoma virus B77	
II	FSV	P140
	Avian sarcoma virus PRCII	P105
III	Avian sarcoma virus Y73	P90
	ESV	P80

This classification is based on homology between transformation-specific genetic sequences[1,2] and gene products[3–5] of the sarcoma viruses. Avian acute leukaemia viruses which also induce sarcomas such as AEV and MC29 are omitted.

0028–0836/81/260675—03$01.00

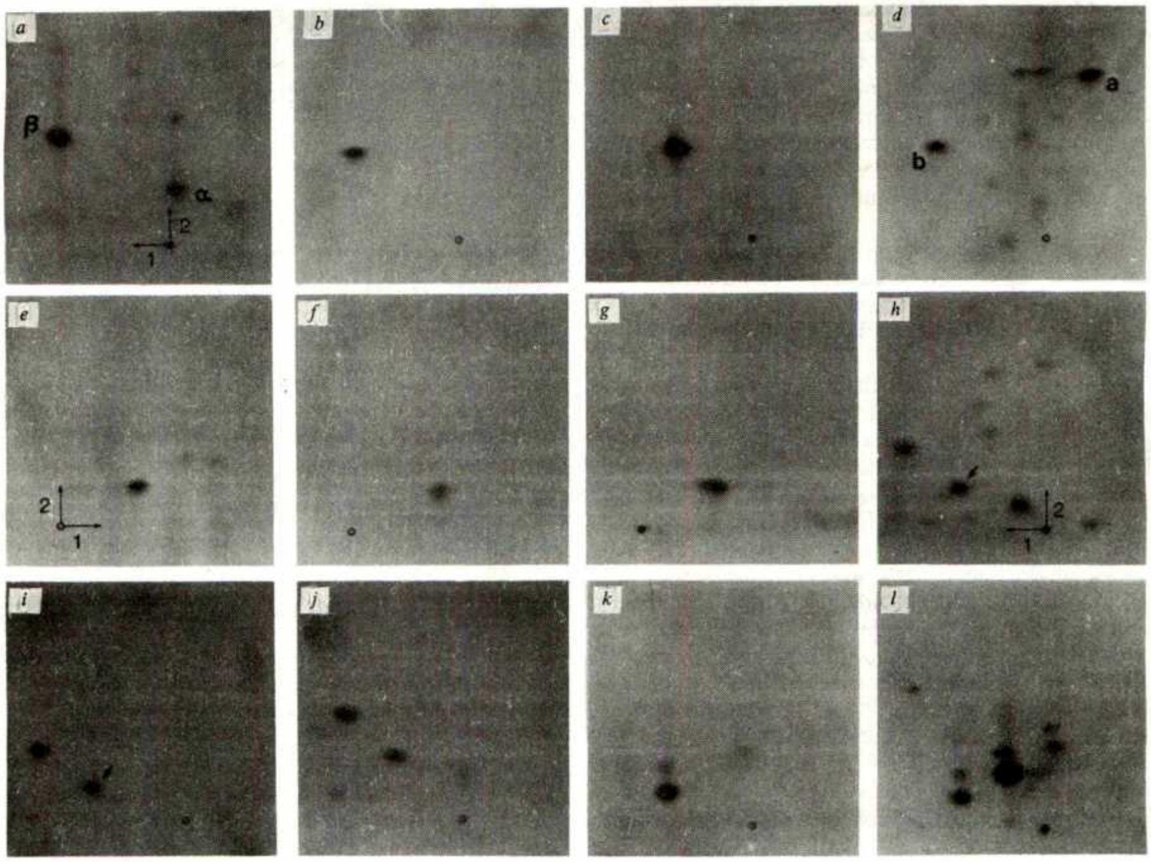


Fig. 1 Tryptic phosphopeptide analysis of RSV pp60^{src}, Y73 P90, ESV P80 and PRCII P105. Proteins were labelled either by metabolic labelling of cells with ^{32}P-orthophosphate in media lacking phosphate or by *in vitro* labelling in immunoprecipitates with [γ-^{32}P]ATP[6,11] (isotopes from ICN). To label pp60^{src}, cells transformed by the Schmidt–Ruppin strain of RSV (SR-RSV-A) were grown overnight in the presence of ^{32}P-orthophosphate. In the case of ESV P80, labelling was for 2 h. As previously described[25], cells were lysed with detergents, and proteins immunoprecipitated with antisera to virion *gag* proteins (for the polyproteins) or antiserum from RSV tumour-bearing rabbits (given by Dr R. Erikson, University of Colorado). To label proteins in *in vitro* phosphotransfer reactions, the viral polyproteins were immunoprecipitated from unlabelled cells and labelled with [γ-^{32}P]ATP (1,000–2,000 Ci mmol^{-1}) at 30 °C for 10 min as described elsewhere[11]. Proteins labelled metabolically ('*in vivo*') or *in vitro* were separated on SDS-polyacrylamide gels. The gels were dried and the protein bands located by autoradiography. The bands were excised, the proteins eluted, precipitated with carier protein, washed in organic solvents, oxidized with performic acid and digested with TPCK-trypsin. This procedure has been described in detail elsewhere[25]. The resultant peptides were separated by thin-layer electrophoresis at 1,000 V for 27 min, in a buffer containing 1% ammonium carbonate *p*H 8.9 (ref. 8). The second-dimension separation was chromatography in *n*-butanol acetic acid/water/pyridine (75:15:60:60) for 4 h. Analyses in *e*, *f* and *g* used, instead, electrophoresis at 600 V for 60 min in a buffer containing *n*-butanol/pyridine/acetic acid/water (2:1:1:80), *p*H 4.5 and chromatography in *n*-butanol/pyridine/acetic acid/water (65:50:10:40) for 4 h. Phosphopeptides were detected by autoradiography on Kodak NS film at room temperature, or Kodak XR film at −70 °C with Du Pont Cronex intensifying screens. The phosphoamino acid content of phosphopeptides α and β of pp60^{src} has been reported by others[8] and confirmed by our analyses (not shown). We have described a similar analysis of ESV P80 and Y73 P90 phosphopeptides *a* and *b* (ref. 5). Phosphopeptides *a* and α contain phosphoserine, whereas *b* and β contain phosphotyrosine. Phosphopeptide α of pp60^{src} is not seen in *e* or *g* as it migrates rapidly off the TLC plate. Phosphoamino acid analyses confirmed that the new V8 cleavage product peptides contained phosphotyrosine (not shown). The maps in *a–g* and *k* are of tryptic digests; the phosphoproteins mapped in *h–j* and *l* were digested with TPCK-trypsin followed by staphyloccoccal V8 protease. The maps are of *a*, pp60^{src} labelled *in vivo*; *b*, ESV P80 labelled *in vitro*; *c*, Y73 P90 labelled *in vitro*; *d*, Y73 P90 labelled *in vivo*; *e*, pp60^{src} labelled *in vivo*; *f*, Y73 P90 labelled *in vitro*; *g*, a mixture of *e* and f; *h*, pp60^{src} labelled *in vivo*; *i*, ESV P80 labelled *in vitro*; *j*, a mixture of *h* and *i*; *k*, PRCII P105 labelled *in vitro*, and *l*, PRCII P105 labelled *in vitro*.

the ability of partially purified pp60^{src} to phosphorylate itself[16,17].

In a previous analysis of the tyrosine phosphorylation sites of PRCII and Y73 (avian sarcoma viruses) and FSV (Fujinami sarcoma virus) proteins[3], we noted that the grouping of these viruses, based on nucleic acid hybridization and peptide mapping, was reflected in the phosphopeptides. We have also shown that the phosphorylation sites on the class III virus proteins are indistinguishable, as might be expected for such closely related viruses[5]. Analysis of pp60^{src} of RSV showed that the tyrosine-containing tryptic peptide of pp60^{src} co-migrated with the equivalent peptides of the Y73 and ESV (Esh sarcoma virus) polyproteins (Fig. 1*a–c*). Like pp60^{src}, Y73 P90 and ESV P80 have additional serine phosphorylation sites that are not found in the helper virus *gag* gene precursor protein Pr76 (Fig. 1*d*). However, the serine phosphorylation sites are not labelled in *in vitro* phosphotransfer reactions (Fig. 1)[5]. Our primary interest is in the *in vitro* kinase target sites, as these are likely to be related to oncogenic function[6,7,14,15]. We analysed the tyrosine-containing tryptic peptides further to discover whether the co-migrating peptides were really identical or showed fortuitous similarity only in the separation conditions of the initial experiment. As a first step we compared electrophoretic migration at a different *p*H, and using a different chromatography buffer (Fig. 1*e–g*). Co-migration was still observed in these conditions, which suggests that the amino acid composition of these phosphotyrosine-containing peptides is very similar.

Separation of the phosphotyrosine-containing tryptic peptides of Y73 P90 and ESV P80 by HPLC on a C18 reverse-phase column shows that the major ^{32}P-labelled peptide of each protein elutes at the same point in the acidic ethanol gradient; the major phosphotyrosine-containing tryptic peptide of pp60^{src} elutes at the same position in these conditions (data not shown, see (ref. 18). Sequential Edman degradation of these ^{32}P-labelled peptides shows that the phosphotyrosine occurs at residue 7 in both Y73 and ESV peptides (Fig. 2), as well as in pp60^{src} (ref. 18).

The DNA sequence of the RSV *src* gene has been determined, and these data have been used to deduce a possible amino acid sequence for pp60^{src} (ref. 19). Partial sequence analysis of the methionine-labelled tryptic peptides of pp60^{src} has demonstrated that the predicted amino acid sequence is correct over most of the length of the protein[18]. Using similar techniques, the amino acid sequence of the phosphotyrosine-

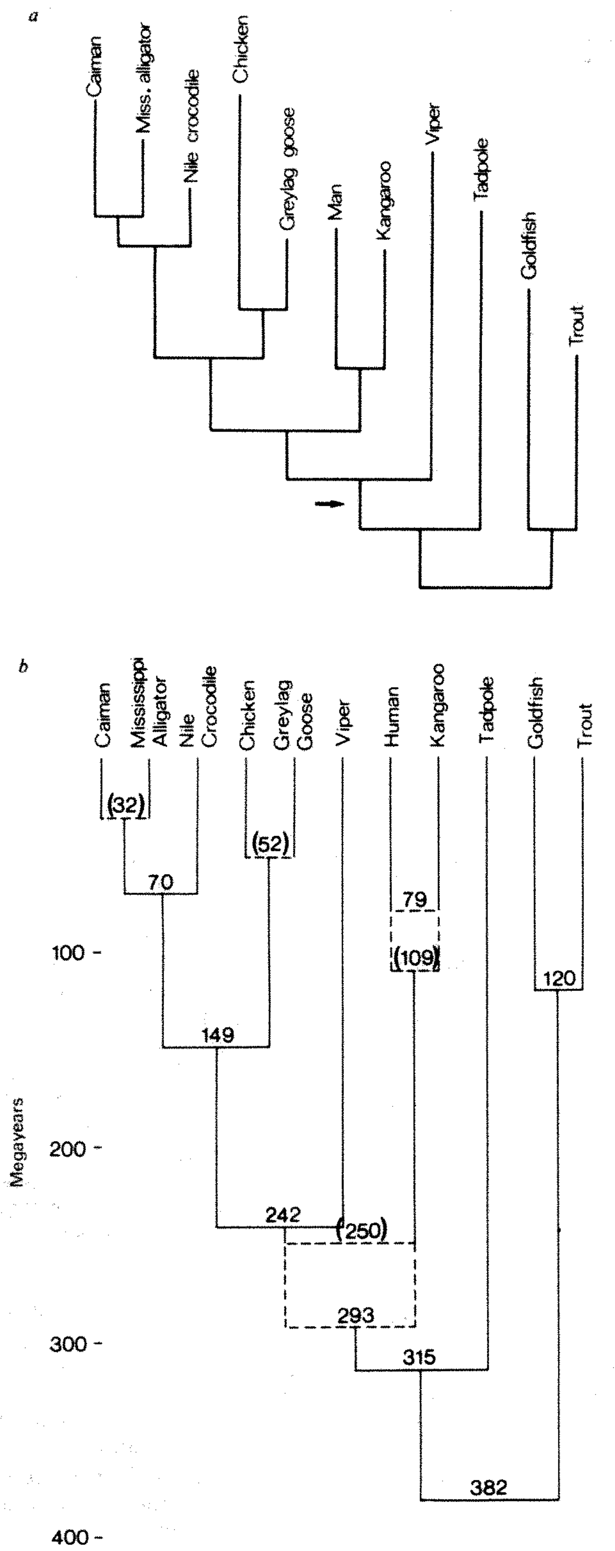


Fig. 2 *a*, Wagner network computed from Table 2 by the method of Farris[22]. The root has been placed midway along the branch separating the two fish from the other species, in accord with conventional zoological opinion. The arrow shows an alternative position of the root at the midpoint between the two species separated by the largest network distance. This network has a total length of 249.5; an alternative network, with the shortest length of 249.0, has a much poorer fit to the observed differences. *b*, Phylogenetic tree based on comparative anatomy and the fossil record. The number at each branching point is an estimate of the minimum date of divergence of the ancestors of the living forms, based on direct fossil evidence. Alternative possible dates, indicated by broken lines and numbers in parentheses, refer to less reliable dates. The justification for some of these dates is given in ref. 19 and the method of dating is further discussed in ref. 23.

pockets and at the subunit contacts, but none of these is unique or could play any part in the allosteric control by bicarbonate ions. All the residues which are essential for the formation of the characteristic T and R structures are either conserved or have been replaced by ones that can serve the same purpose equally well.

Do the crocodilian sequences tell us anything interesting about the phylogenetic relationships between these reptiles and other groups of bony vertebrates? The α-globin sequences of 21 mammals, 3 birds, 3 fish, 1 amphibian and 1 other reptile (viper) are now known, but not the β-globin sequence of the viper. We have selected two species from among each of the mammals, birds and fish and compared their α-globin sequences with those of the crocodilians, the bullfrog tadpole and the viper. Table 2 shows a matrix of amino acid differences between all possible pairs of sequences; the data in Fig. 2*a* have been calculated from this matrix by an algorithm that does not allow the difference between any pair of sequences to be less than the observed one; it attempts to minimize the total length of the network and to maximize its fit to the observed differences. It shows that the crocodilian sequences are most similar to those of the birds, consistent with fossil evidence suggesting that birds arose from archosaurian reptiles. As might be expected, the number of amino acid differences and the network distances between each of the crocodilians and the viper (a lepidosaur reptile) are much larger than between the crocodilians and the birds, but it is surprising that they are also much larger than between the crocodilians and the mammals. Figure 2*b* shows, for comparison, a zoologically conventional phylogenetic tree, derived from the evidence of comparative anatomy and the fossil record, which shows the crocodilians as more closely related to the viper than the mammals. For the rest, Fig. 2*a* and *b* are similar in pattern, although Fig. 2*b* has very unequal pairs of branches arising from the same node. These support the conclusions of Romero-Herrera *et al.*[18,19], based on myoglobin, that the rates of change differ on different branches of the phylogenetic tree. Without knowing more reptile sequences we cannot be sure, therefore, whether the large differences between the crocodilians and the viper suggest a divergence between the ancestors of archosaurian and lepidosaurian reptiles earlier than that between the archosaurian reptiles and mammals, or are due merely to accelerated fixation of mutations along one particular evolutionary lineage. This question is worth following up.

We thank Mr A. Stangl and Mrs B. Schrank for their help in determining the amino acid sequences, Professors K. Johansen and R. Coulson for gifts of crocodile blood, Drs S. E. V. Phillips and B. Shaanan for their help with Fig. 1 and Professor Howard Morris for checking an N-terminal sequence by mass spectrometry.

Received 24 March; accepted 8 May 1981.

1. Baldwin, J. M. *Prog. Biophys. molec. Biol.* **29**, 225–320 (1975).
2. Perutz, M. F. *Br. med. Bull.* **32**, 191–208 (1976).
3. Kilmartin, J. V. *et al. Biochim. biophys. Acta* **534**, 15–25 (1978).
4. Kilmartin, J. V. & Rossi-Bernardi, L. *Physiol. Rev.* **53**, 836–890 (1973).
5. Kilmartin, J. V. *Trends biochem. Sci.* **2**, 247–250 (1977).
6. Bauer, C. *et al. J. biol. Chem.* (in the press).
7. Bauer, C. & Jelkmann, W. *Nature* **269**, 825–827 (1977).
8. Leclercq, F., Schnek, A. G., Braunitzer, G., Stangl, A. & Schrank, B. *Hoppe-Seyler's Z. physiol. Chem.* (in the press).
9. Arnone, A. *Nature* **237**, 146–149 (1972).
10. Bunn, H. F. & Briehl, R. W. *J. clin. Invest.* **49**, 1088–1095 (1970).
11. Arnone, A., Rogers, P. H. & Briley, P. D. in *Biophysics and Physiology of Carbon Dioxide* (eds Bauer, C., Gros, G. & Bartels, H.) 67–74 (Springer, Berlin, 1980).
12. O'Donnell, S., Mandaro, R., Schuster, T. M. & Arnone, A. *J. biol. Chem.* **254**, 12204–12208 (1979).
13. Imai, K. *J. biol. Chem.* **249**, 7607–7612 (1974).
14. Perutz, M. F., Fersht, A. R., Simon, S. R. & Roberts, G. C. K. *Biochemistry* **13**, 2174–2186 (1974).
15. Fermi, G. & Perutz, M. F. *J. molec. Biol.* **114**, 421–431 (1977).
16. Perutz, M. F. *et al. J. molec. Biol.* **138**, 649–670 (1980).
17. Jelkmann, W. & Bauer, C. *Comp. Biochem. Physiol.* **A65**, 331–336 (1980).
18. Romero-Herrera, A. E., Lehmann, H., Joysey, K. A. & Friday, A. E. *Nature* **246**, 389–395 (1973).
19. Romero-Herrera, A. E., Lehmann, H., Joysey, K. A. & Friday, A. E. *Phil. Trans. R. Soc.* **B283**, 61–163 (1978).
20. Fermi, G. *J. molec. Biol.* **97**, 237–256 (1975).
21. R. Diamond in *Biomolecular Structure, Conformation and Evolution* Vol. 1 (ed. Srinivasan, R.) 567–588 (Pergamon, Oxford, 1980).
22. Farris, J. S. *Am. Nat.* **106**, 645–668 (1972).
23. Joysey, K. A. *Symp. zool. Soc. Lond.* **46** (in the press).

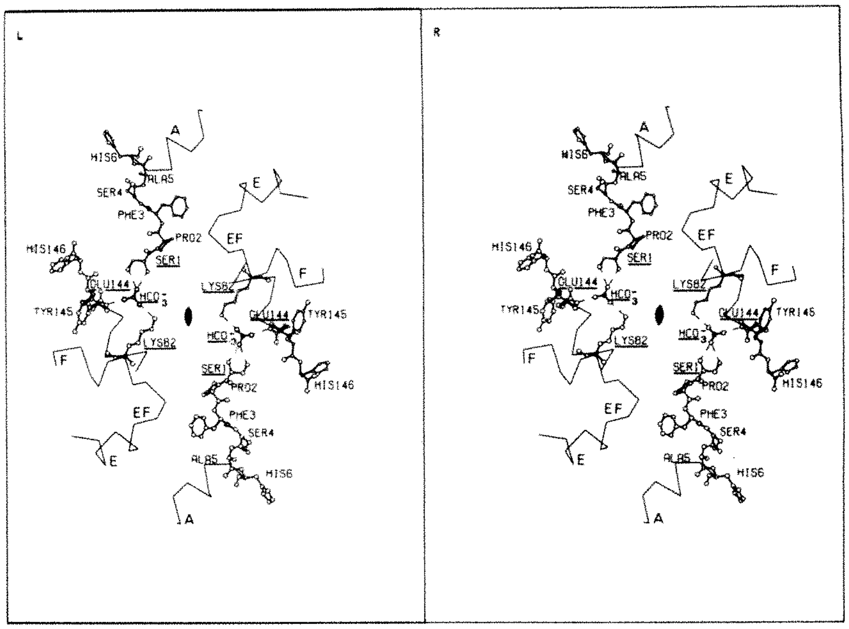


Fig. 1 Stereo drawing of proposed bicarbonate binding site between the two β-chains of caiman deoxyhaemoglobin. The central sign marks the dyad symmetry axis. Bicarbonates and their binding residues are underlined. Capital letters mark helical and inter-helical segments. The drawing was prepared using the PLUTO program and Fermi's refined atomic coordinates of human deoxyhaemoglobin[20]. Certain of the amino acid residues in the atomic model of this haemoglobin were replaced by those in caiman; the bicarbonate ions were attached to them and all the new atomic positions were then measured manually. These rough coordinates were then corrected to standard bond distances and angles using Diamond's BILDER program[21] on the Evans and Sutherland Picture System I linked to a DEC PDP 11/50 computer. The hydrogen bond distances quoted in the text were derived from this computer-generated model. The hydrogen bond from HCO_3^- to COO^- of Glu 144 looks straight in the picture, but in fact the two groups lie at different heights, making the bond angle COH 110°, as it should be.

oxygen can form a salt bridge with Lys EF6(82) and the third oxygen can donate a hydrogen bond to one of the carboxylate oxygens of Glu HC1(144). In this model all hydrogen bonds come out reasonably: Ser $N–O_1 = 2.8$ Å; Ser $O_\gamma–O_1 = 3.3$ Å; Lys $N_\zeta–O_2 = 2.9$ Å; Glu $O_\epsilon–O_3 = 2.7$ Å. All side chains are in the strain-free, straight, staggered conformation. Even so, the model should be checked by X-ray analysis. In the haemoglobins of the Nile crocodile and the Mississipi alligator, where the N-terminal sequence is acetyl-Ala-Ser-Phe, the α-NH of acetylalanine could donate a hydrogen bond to O_1 of the bircarbonate at the same distance as the α-NH_3^+ does in caiman, provided the chain turns a corner at Ser 2 at the same angle as it does at Pro 2 in caiman. Again, this should be checked by X-ray analysis.

The salt bridges and hydrogen bonds made by the two symmetry-related bicarbonate ions would stabilize the relative positions taken up by the two β-chains in the quaternary T structure; furthermore, by immobilizing Glu HC1(144), they would help to clamp the three C-terminal residues Glu-Tyr-His in their rigid deoxy conformation, with the tyrosine hydrogen-bonded to Val FG5β and the histidine salt-bridged to Asp FG1β and Lys C5α. On transition to the R structure, the two β-chains move relative to one another by several Ångstroms and their C-terminal residues become mobile, so that the bicarbonate binding sites are destroyed.

Other small anions also lower the oxygen affinity, but much less than bicarbonate. When the concentrations of chloride, nitrate, perchlorate or phosphate are increased from 0.01 M to 0.5 M at *p*H 7.0 and 25 °C, P_{50} is lowered by factors of 2.0, 1.6, 1.7 and 1.3, respectively, whereas the same increase in the concentration of bicarbonate lowers P_{50} 13-fold[6]. In human haemoglobin, chloride probably binds to Lys 82β (ref. 16). If the same is true of caiman haemoglobin, the binding of chloride and bicarbonate should be competitive; in fact, a rise in $[Cl^-]$ from 0.01 to 0.5 M approximately halves the effect of HCO_3^- on P_{50}. NO_3^- is a planar ion like HCO_3^-, and the NO distances are only 0.1 Å shorter than those of CO. NO_3^- could therefore bind to Ser 1 and Lys 82, but it has no hydrogen to donate to Glu 144; nor does ClO_4^-. HPO_4^{2-} could do so, but because this ion is tetrahedral, binding of one oxygen to Ser 1 and another to Lys 82 makes the OH bond of the third oxygen point away from the carboxylate of Glu 144, so that no hydrogen bond can be formed.

The decrease in oxygen affinity brought about by the interaction of crocodilian haemoglobin with bicarbonate ensures that oxygen is released from the blood to the tissues at a relatively high partial pressure of oxygen. If the haemoglobin were insensitive to bicarbonate, the venous P_{O_2} would be only 7 torr; interaction with bicarbonate raises this to 27 torr, thus creating a large enough pressure head for the flow of oxygen from the blood to the tissues[17]. In human haemoglobin, the same relative rise in P_{O_2}, from 19 to 40 torr, can be brought about only by the combined effects of CO_2 and diphosphoglycerate. It is surprising that the simple and direct reciprocating action between oxygen and one of the end products of oxidative metabolism has not been adopted by other vertebrates. We do not know what advantage it gives to the crocodilians.

Our results show that an entirely new function can evolve in a protein by no more than three amino acid substitutions, requiring only four nucleotide base changes; just two more amino acid substitutions, or three base changes, are needed for inhibition of the old functions (oxygen-linked phosphate, carbamino CO_2^- and Cl^- binding). Most of the other ~100 substitutions that distinguish crocodilian from human haemoglobins are conservative and would have little if any effect on the oxygen equilibrium. There are some significant substitutions in the haem

Table 2 Number of amino acid substitutions between the α-globin chains of different species

	Man	Caiman	Mississippi alligator	Nile crocodile	Goldfish	Trout	Tadpole	Chicken	Goose	Viper
1. Caiman	46									
2. Missippi alligator	47	20								
3. Nile crocodile	44	21	17							
4. Goldfish	65	68	68	64						
5. Trout	60	64	64	66	43					
6. Tadpole	61	65	67	66	70	63				
7. Chicken	35	46	49	45	72	70	70			
8. Goose	41	39	35	30	64	61	64	30		
9. Viper	50	64	66	62	76	70	69	57	59	
10. Kangaroo	27	43	43	43	67	60	60	41	39	54

1, *Caiman crocodilus*; 2, *Alligator mississippiensis*; 3, *Crocodylus niloticus*; 4, *Carassius auratus*; 5, *Salmo irideus*; 6, *Rana catesbeiana*; 8, *Anser anser*; 9, *Vipera aspis*; 10, *Macropus giganteus*.

Allosteric regulation of crocodilian haemoglobin

M. F. Perutz*, C. Bauer†, G. Gros‡, F. Leclercq§, C. Vandecasserie§, A. G. Schnek§, G. Braunitzer‖, A. E. Friday¶ & K. A. Joysey¶

* MRC Laboratory of Molecular Biology, Cambridge CB2 2QH, UK
† Institut für Physiologie, Universität, 8400 Regensburg, FRG
‡ Physiologisches Institut, Universitätsklinikum 4300 Essen, FRG
§ Chimie General I, Université Libre, Bruxelles, Belgium
‖ Max-Planck Institut für Biochemie, 8033 Martinsried, FRG
¶ University Museum of Zoology, Cambridge CB2 3EJ, UK

The oxygen affinity of most vertebrate haemoglobins in the absence of diffusible electrolytes is much higher than that of blood. In the red cell this affinity is lowered by organic phosphates, hydrogen ions, chloride ions and CO_2 (refs 1–5). Similarly, crocodilian haemoglobin also has a much higher oxygen affinity than crocodile blood, but this is due to bicarbonate ions[6], as neither phosphates nor carbamino CO_2^- lower its oxygen affinity, and chloride does so only weakly[7]. The complete sequences of the haemoglobins of the caiman, the Nile crocodile and the Mississippi alligator (to be reported elsewhere[8]) show 102 substitutions between human and caiman, and 123 between human and the other two crocodilian haemoglobins. Here we consider how these substitutions may explain the changes in allosteric control, and also their bearing on the phylogenetic relationships between the crocodilians and other groups of bony vertebrates. We propose that a few of the substitutions abolish or weaken the binding sites for the usual allosteric effectors and create a new pair of binding sites which are complementary to bicarbonate ions in the deoxy (T) structure, but not in the oxy (R) structure.

Consider first the loss of allosteric inhibition by organic phosphates, carbamino CO_2^- and chloride (Table 1). In human deoxyhaemoglobin, organic phosphates are bound by the four pairs of cationic residues that lie in the cavity between the two β-chains: Val NA1 (1), His NA2 (2), Lys EF6 (82) and His H21 (143)[9]. In the three crocodilian haemoglibins, the histidines are replaced by neutral residues. In two of the haemoglobins the α-amino groups are acetylated; in the third, a proline in position 2 forces a bend in the chain which places the α-amino groups out of reach of the phosphates. This leaves only the two lysines which are known from human fetal haemoglobins F_I to be insufficient for phosphate binding, although His NA2 is still present there[10]. In human deoxyhaemoglobin, carbamino CO_2^- bound to Val 1α accepts a hydrogen bond from Ser H14 (131) α (ref. 11). In the three crocodilian haemoglobins this is replaced by Ala. This replacement may also inactivate one of the strong chloride binding sites which probably coincides with the CO_2^- site[12]. In human deoxyhaemoglobin, another pair of CO_2 molecules competes with organic phosphates for binding to Val 1β; when bound, each carbamino CO_2^- is stabilized by a salt bridge to Lys 82β (ref. 11). In caiman haemoglobin, the bend in the chain forced by Pro 2 inhibits that bridge; in the other two crocodilian haemoglobins, the blocking of Val 1β inhibits binding of CO_2.

Two equivalents of bicarbonate ion per tetramer are bound to caiman deoxyhaemoglobin with an association constant of $2.0\times10^{-3}\ M^{-1}$, whereas oxyhaemoglobin shows no significant binding. In the presence of CO_2, discharge of oxygen produces no change in *p*H, because the protons taken up by haemoglobin are neutralized by those set free in the formation and binding of bicarbonate[6]. At *p*H 7.4 and 20 °C in 0.1 M NaCl, application to 40 torr CO_2 raises P_{50} ninefold from 3.4 to 30.0 torr; this is comparable to the 10-fold rise, from 4.35 to 44.0 torr, produced by the addition of 2 mM inositolhexaphosphate (IHP) to a similar solution of human haemoglobin A at 25 °C (ref. 13). In human haemoglobin A, IHP reduces Hill's coefficient from 2.9 to 2.4; in crocodile haemoglobin, CO_2 reduces it to the same value. In both cases the allosteric equilibrium is shifted too far towards the T structure for oxygen binding to be fully cooperative. At *p*H 7.0, addition of CO_2 to caiman methaemoglobin produces a large negative circular dichroism peak at 287 nm, diagnostic of the quaternary T structure; the same effect is produced by IHP in human methaemoglobin A[14,15].

We have examined all possible sites in the model of human deoxyhaemoglobin where the amino acid replacements indicated by the three crocodilian sequences could possibly cause oxygen-linked binding of bicarbonate ions, and have found only one pair of sites with the right stereochemistry. These lie in the cavity between the two β-chains where organic phosphates or carbamino CO_2^- are bound in other species, and are formed by Lys EF6(82) and Glu HC1(144) of one β-chain together with the N-terminal residue of its partner chain (Fig. 1, Table 1). In caiman haemoglobin, the N-terminal sequence is Ser-Pro-Phe-. A model of the bicarbonate binding site was first built manually by replacing the side chains in human deoxyhaemoglobin with those of caiman, and was then refined by computer. The main chain from the α-carbon of Pro 2 onwards was held in the same position as in human deoxyhaemoglobin. At this atom the chain is forced to turn a corner at an angle defined by the stereochemistry of the proline. This rigid turn brings the N-terminal serine within exact reach of the bicarbonate ion, so that one of the bicarbonate oxygens can form a salt bridge with the α-NH_3^+ and can also accept a rather long hydrogen bond from the serine OH. The second bicarbonate

Table 1 Amino acid replacements for allosteric control in crocodile haemoglobin

Position in			Species					
Structure		Sequence	Human	Other bony vertebrates	Caiman	Nile crocodile	Mississippi crocodile	Functions
NA1	β	1	Val	Val or Gly	Ser	Ac-Ala	Ac-Ala	In human and many other vertebrate deoxy-Hbs, the α-amino groups bind either organic phosphate or carbamino CO_2^-, which in turn forms a salt bridge with Lys 82 of the same β-chain. In crocodile Hb, blocking of α-NH_3^+ or Pro in position 2 inhibits these functions
NA2	β	2	His	Gln, Asn, Glu, His, Asp, Met	Pro	Ser	Ser	In human Hb, His, and in many other vertebrates Gln or Asn, bind organic phosphates
EF6	β	82	Lys	Lys	Lys	Lys	Lys	Lys binds diffusible anions in the internal cavity of deoxy-Hbs. Invariant in all vertebrate Hbs except trout I which lacks all oxygen-linked functions and where it is replaced by Leu*
H21	β	143	His	Lys, Arg, Ser, His	Ala	Ala	Ala	His, Lys or Arg form salt bridges with organic phosphates
HC1	β	144	Lys	Arg, Ala, Ser, Gln, Lys	Glu	Glu	Glu	In human Hb, Lys is free and the same is probably true of Arg, Ser or Gln in other species. Glu can accept an H-bond from HCO_3^- whose charge is compensated by Lys 82
H14	α	131	Ser	Very variable	Ala	Ala	Ala	In human Hb, OH of Ser forms an H-bond with carbamino CO_2^- or Cl^- bound to Val 1α, but this bond is absent in many species

* Unpublished data (D. Barra, F. Bossa and M. Brunori).

0028-0836/81/260682—03$01.00

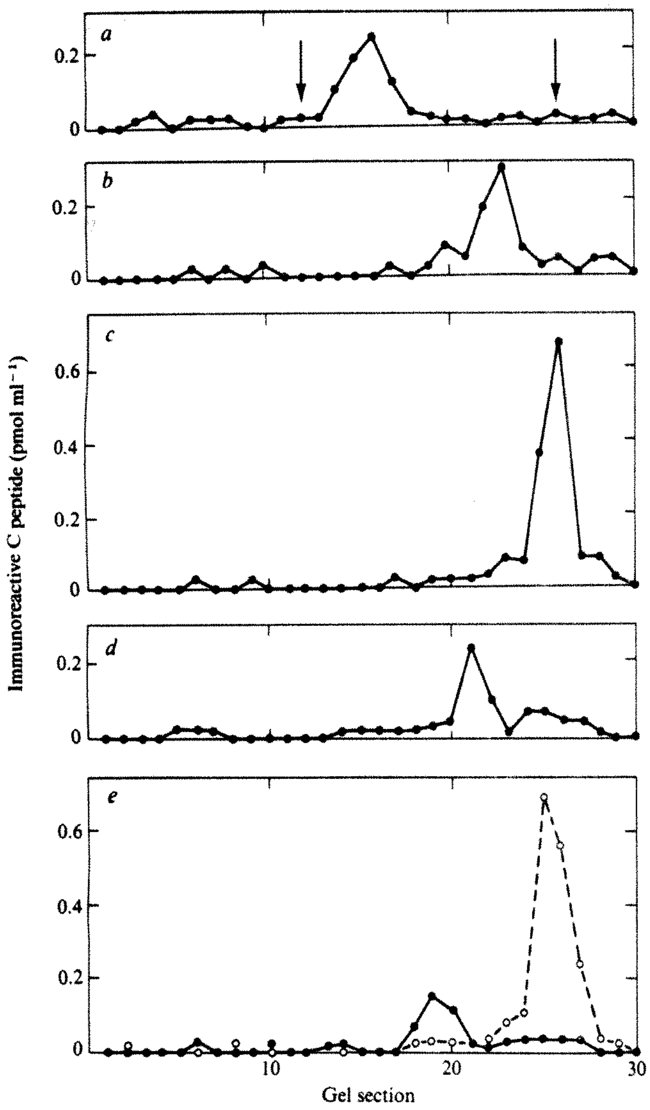


Fig. 2 Polyacrylamide gel electrophoresis of purified proinsulin-like material before and after enzymatic digestion and acetylation. Enzymatic digestions proceeded as described for Fig. 1 except that the buffer was 0.1 M *N*-ethylmorpholine adjusted to *p*H 8 with acetic acid; after digestion the buffer and solvent were removed under vacuum and the residue was dissolved in 0.1 ml of electrophoresis buffer before electrophoresis on 15% (w/v) acrylamide gels at *p*H 8.7 (ref. 21). Peptides from 1.5-mm sections of the gel were eluted into 0.5 ml of buffer[18] for 20 h at 4 °C and aliquots of the resulting solutions were subjected to the C-peptide radioimmunoassay[18]. *a*, Native peptide; *b*, peptide digested by trypsin; *c*, peptide sequentially digested by trypsin and carboxypeptidase B; *d*, peptide sequentially digested by carboxypeptidase B and trypsin; *e*, peptide treated with acetic anhydride [four 5-μl aliquots over a 2-h period at 22 °C in 1 ml of 2 M sodium acetate (solid line) or in 1 ml of 2 M ammonium acetate (dashed line)] before dialysis against 3 M acetic acid, evaporation and sequential digestion by trypsin and carboxypeptidase B. Approximately 2 pmol of peptide was used in each case. The left and right arrows in panel *a* show the positions taken by native human proinsulin and human C peptide, respectively. The tracking dye (bromophenol blue) appeared in gel section 30.

for a product having the C peptide joined to the A chain of desnonapeptide insulin), but to a position well behind that taken by authentic human C peptide. Control studies showed that (1) the intermediate treated with acetic anhydride in the presence of ammonium ion (to prevent amino group modification) and then subjected to purification and analysis yielded the native C peptide (Fig. 2*e*, dashed line), and (2) enzymatic digestion of acetylated human proinsulin yielded a peptide (ε-*N*-acetyl-C-peptidyllysine) having the electrophoretic mobility of native human C peptide (not shown).

The above results demonstrate both the presence of Lys 64 in the region joining the C peptide to the insulin A chain in the proinsulin intermediate and the presence of a nearby amino acid alteration. Although mutations resulting in a substitution or deletion at Arg 65 [yielding the connecting region sequence Gln-Lys-(X)-Gly] or resulting in a proline for glycine substitution at position A1 (yielding the sequence Gln-Lys-Arg-Pro) could both account for this sensitivity, the latter possibility can be excluded: tryptic digestion of an intermediate with the latter sequence would have resulted in arginyl desoctapeptide insulin rather than the desoctapeptide insulin identified on the gel of Fig. 2*b*. We thus conclude that the biosynthetic intermediate results from failure of cellular enzymes to process correctly a proinsulin variant having an amino acid alteration at Arg 65. The relatedness of this intermediate to an uncleaved proalbumin variant containing a glutamine for arginine substitution[22] emphasizes the essential pairing of basic residues for substrate recognition of both peptide and protein precursors by processing enzymes.

Our analysis of the variant proinsulin intermediate des-Arg 31, Arg 32-[X 65] proinsulin in patients with familial hyperproinsulinaemia identifies a new locus of mutation within the insulin gene. These loci include sites in the B-chain region[1–4], the connecting sequence between the C peptide and the insulin A chain (this report), the 3′-untranslated region[5,6] and the two intervening sequences[6]. Nevertheless, the structural and physiological consequences of these changes vary considerably: the first results in a normally processed, but functionally abnormal hormone; the second results in an abnormally processed intermediate of conversion (with the likely potential for conversion to the normal hormone); the third and fourth probably result in completely normal hormone products. Another proinsulin intermediate, previously isolated from a different family with hyperproinsulinaemia, and reported to have the C peptide joined to the insulin B chain[8], has not been analysed in detail. The complexity of both post-transcriptional[5,6] and post-translational[11–13,23,24] processing in insulin biosynthesis suggests still other sites (especially those for excision and ligation of precursor mRNA and for cleavage of the signal peptide from preproinsulin) at which mutations would yield abnormal gene products.

This work was supported by grants AM 18347, AM 13941 and AM 20595 from the National Institute of Arthritis, Metabolism, and Digestive Diseases of the NIH. D.C.R. is the recipient of National Science Research Award AM 05723 from the NIH. H.S.T is the recipient of Research Career Development Award AM 00145 from the NIH.

Received 12 January; accepted 17 March 1981.

1. Tager, H. *et al. Nature* **281,** 122–125 (1979).
2. Given, B. D. *et al. New Engl. J. Med.* **302,** 129–135 (1980).
3. Olefsky, J. M., Saekow, M., Tager, H. & Rubenstein, A. H. *J. biol. Chem.* **255,** 6098–6105 (1980).
4. Tager, H. *et al. Proc. natn. Acad. Sci. U.S.A.* **77,** 2410–2414 (1980).
5. Bell, G. I. *Nature* **284,** 26–32 (1980).
6. Ullrich, A., Dull, T. J., Gray, A., Brosius, J. & Sires, I. *Science* **209,** 612–615 (1980).
7. Gabbay, K. H., DeLuca, K., Fischer, J. N., Mako, M. E. & Rubenstein, A. H. *New Engl. J. Med.* **294,** 911–915 (1976).
8. Gabbay, K. H., Bergenstal, R. M., Wolff, J., Mako, M. E. & Rubenstein, A. H. *Proc. natn. Acad. Sci. U.S.A.* **76,** 2881–2885 (1979).
9. Kanazawa, Y., Hayashi, M., Ikeuchi, M., Hiramatsu, K. & Kosaka, K. *Eur. J. clin. Invest.* **8,** 327–331 (1978).
10. Kanazawa, Y. *et al. Proc. Symp. Proinsulin, Insulin and C-Peptide,* Tokoshima (eds Baba, S., Kaneko, T. & Yanihara, N.) 262-269 (Excerpta Medica, Amsterdam, 1979).
11. Steiner, D. F. *Recent Prog. Horm. Res.* **24,** 207–282 (1969).
12. Chance, R. E. in *Diabetes* (eds Rodrigues, R. R. & Vallance-Owen, J.) 292–305 (Excerpta Medica, Amsterdam, 1971).
13. Steiner, D. F. *Diabetes* **26,** 322–340 (1977).
14. Horwitz, D. L., Starr, J. I., Mako, M. E., Blachard, W. G. & Rubenstein, A. H. *J. clin. Invest.* **55,** 1278–1283 (1975).
15. Cuatrecasas, P. *J. biol. Chem.* **245,** 3059–3065 (1970).
16. Starr, J. I., Horwitz, D. L., Rubenstein, A. H. & Mako, M. E. *Methods of Hormone Radioimmunoassay* (eds Jaffe, B. M. & Behrman, H. L.) 613–629 (Academic, New York, 1979).
17. Varandani, P. T. *Biochemistry* **6,** 100–104 (1967).
18. Kuzuya, H., Blix, P. M., Horwitz, D. L., Steiner, D. F. & Rubenstein, A. H. *Diabetes* **26,** 22–29 (1977).
19. Faber, O. K. *et al. Diabetes* **27,** Suppl. 1, 170–177 (1978).
20. Cole, R. D. *Meth. Enzym.* **11,** 206–208 (1967).
21. Davis, B. J. *Ann. N. Y. Acad. Sci.* **121,** 404–422 (1964).
22. Brennan, S. O. & Carrell, R. W. *Nature* **274,** 908–909 (1978).
23. Chan, S. J., Keim, P. & Steiner, D. F. *Proc. natn. Acad. Sci. U.S.A.* **73,** 1964–1968 (1976).
24. Tager, H. S. *et al. Ann. N. Y. Acad. Sci.* **343,** 133–147 (1980).

Sera from the three affected members of the family (a 62-yr-old man with glucose intolerance, his 57-yr-old sister and his 42-yr-old son) all showed marked elevations in fasting concentrations of immunoreactive insulin (70–100 μU ml^{-1} compared with normal values of <25 μU ml^{-1}) and abnormally high ratios of proinsulin-like material to insulin (9–10 compared with normal values of <0.25, ref 14). Insulin-related components from the sera of the two younger patients were purified by immunoaffinity chromatography[2,15] using guinea pig anti-porcine insulin serum pooled from several immunized animals. Gel filtration of the eluted peptides on Bio-Gel P-30 (equilibrated with 3 M acetic acid containing 5 μg ml^{-1} of bovine serum albumin) resulted in the separation of the 9,000-molecular weight (MW) proinsulin-like peptide from the smaller amount of 6,000-MW insulin. In all, 400 ml of serum was processed in this manner, yielding ~150 pmol of the proinsulin-like component; recoveries averaged 90%.

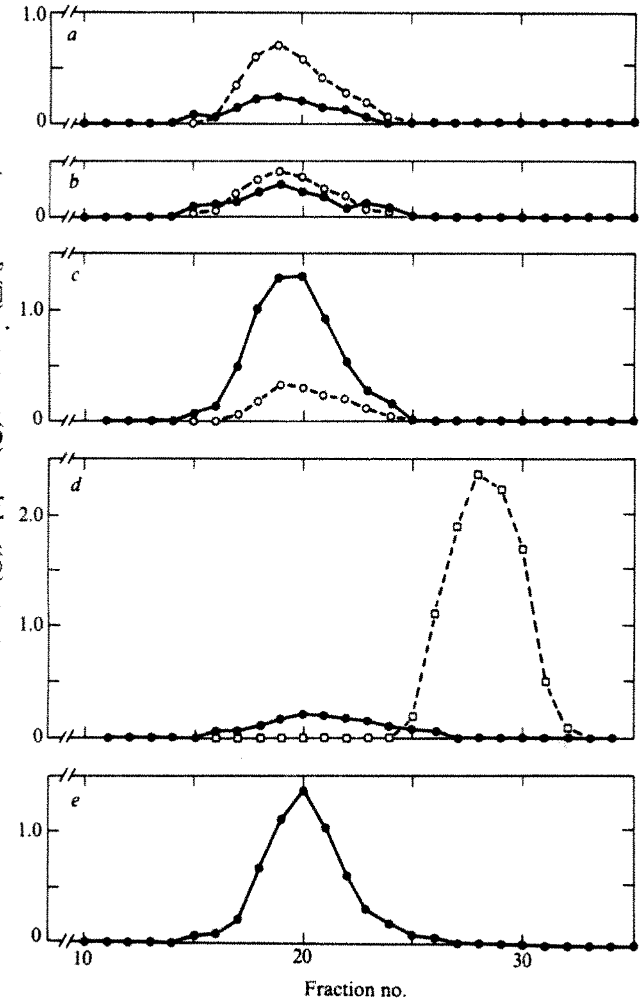


Fig. 1 Gel filtration and immunoreactivity of purified proinsulin-like material before and after oxidative sulphitolysis. The column of Bio-Gel P-10 (0.9×90 cm) was equilibrated with a buffer containing 0.1 M Tris brought to *p*H 7.6 with HCl, 0.05 M NaCl, 0.2 mg ml^{-1} sodium azide and 5 μg ml^{-1} bovine serum albumin; 1.2-ml fractions were collected. Aliquots of column fractions were assayed for immunoreactive C peptide using antiserum M1230[18] (●), insulin using antiserum GP-I[16] (○) and B chain using an antiserum donated by P. R. Varandani[17] (□). Human C peptide, porcine insulin and porcine bis-*S*-sulpho B chain served as the respective immunoassay standards. *a*, Native peptide, untreated fractions; *b*, native peptide, trypsin-digested fractions [100 μg ml^{-1} tosylphenylalanine chloromethyl ketone-treated trypsin (Worthington), 10 mM $CaCl_2$, 60 min, 37° C, briefly heated to 100° C after digestion]; *c*, native peptide, fractions sequentially digested by trypsin and carboxypeptidase B [trypsin digestion as above, 100 μg ml^{-1} carboxypeptidase B (Boehringer-Mannheim), 60 min, 37 °C, briefly heated to 100 °C after digestion]; *d*, oxidatively sulphitolysed peptide [1 ml of 0.2 M sodium phosphate buffer at *p*H 7.4 containing 8 M urea, 1.75% (w/v) sodium tetrathionate and 3.5% (w/v) sodium bisulphite, 3 h, 37 °C], untreated column fractions; *e*, oxidatively sulphitolysed peptide, fractions sequentially digested by trypsin and carboxypeptidase B. Approximately 7 pmol of peptide was gel-filtered in each case. Native proinsulin and authentic bis-*S*-sulpho B chain appeared in fractions 19 and 28, respectively, during standardization of the column.

Because of the small amounts of proinsulin-like material obtainable from serum, our approach to examining the structure of the 9,000-MW peptide relied on sensitive radioimmunoassays for insulin[16], insulin B chain[17] and proinsulin C peptide[18] to detect picomole quantities of material and on specific enzymatic and chemical manipulations to degrade the peptide to known components. Figure 1*a* shows that the purified and gel-filtered 9,000-MW peptide reacts, on a molar basis, only one-third as well with C-peptide antibodies as it does with insulin antibodies. Treatment of the column fractions with trypsin slightly enhanced the C-peptide immunoreactivity and diminished the insulin immunoreactivity (Fig. 1*b*). Sequential treatment with trypsin and carboxypeptidase B, however, resulted in a fivefold enhancement of C-peptide immunoreactivity (Fig. 1*c*). The known sensitivity of our C-peptide antiserum to the COOH-terminal sequence of human C peptide[19] and the specific effect of carboxypeptidase B suggest a masking of the COOH-terminal region of the C peptide within the 9,000-MW form.

To determine whether the 9,000-MW peptide was a two-chained intermediate of conversion or single-chained proinsulin, the peptide was subjected to oxidative sulphitolysis, a procedure which irreversibly breaks both interchain and intrachain disulphide bonds[20]. As shown in the gel filtration profile of Fig. 1*d*, the oxidized peptide separated into two distinct components, a higher-molecular-weight peptide exhibiting C-peptide immunoreactivity and a lower-molecular-weight peptide co-eluting with native B chain and showing B-chain immunoreactivity. Although treatment of individual column fractions with trypsin slightly augmented the C-peptide immunoreactivity and obliterated B-chain immunoreactivity (not shown), sequential treatment with trypsin and carboxypeptidase B, as before, enhanced the higher-molecular-weight C-peptide immunoreactivity fivefold (Fig. 1*e*). The experiments illustrated in Fig. 1 identify the patients' proinsulin-like peptide as a two-chained intermediate of conversion in which the C peptide remains joined to the insulin A chain (des-Arg 31, Arg 32-proinsulin).

We examined the proinsulin intermediate and the products of its enzymatic digestion in greater detail by polyacrylamide gel electrophoresis at *p*H 8.7 (ref. 21). As shown in Fig. 2*a*, the anodal mobility of the native peptide was greater than that of human proinsulin, a result consistent with the excision of basic amino acid residues from proinsulin during formation of the conversion intermediate. Digestion of the peptide with trypsin resulted in a more rapidly migrating fragment with C-peptide immunoreactivity (Fig. 2*b*). However, only sequential treatment of the intermediate with trypsin and carboxypeptidase B resulted in a fragment co-migrating with native human C peptide and showing a major enhancement in C-peptide immunoreactivity (Fig. 2*c*). As indicated by Fig. 2*d*, reversing the order of digestion with trypsin and carboxypeptidase B failed to yield native human C peptide. Analysis of the gels of Fig. 2*a*–*d* for insulin-related peptides showed that the insulin and C-peptide immunoreactivities of the undigested 9,000-MW peptide migrated to the same position, but that treatment with trypsin alone and with trypsin plus carboxypeptidase B yielded peptides co-migrating with desoctapeptide insulin (lacking residues B23–B30) and desnonapeptide insulin (lacking residues B22–B30), respectively (data not shown).

Further analysis of the proinsulin intermediate for an amino acid substitution in its C peptide–A chain connecting region proceeded from selective protection of amino groups by acetylation. We reasoned that acetylation would not affect tryptic cleavage of the C peptide from an intermediate lacking Lys 64, whereas acetylation would prevent removal of the C peptide from an intermediate lacking Arg 65. As shown in Fig. 2*e* (solid line), polyacrylamide gel electrophoresis of the trypsin- and carboxypeptidase B-digested, acetylated intermediate yielded a product migrating further than the native peptide (as expected

Sanger *et al.*[8]. As the coding Lb sequence contains a *Hin*dIII cleavage site in the position corresponding to amino acids 64–65, the 5′ ends resulting from a *Hin*dIII digestion of the 7.5-kilobase fragment could be labelled with ^{32}P using [γ-^{32}P]ATP and T_4 polynucleotide kinase. Appropriate DNA fragments were then subjected to DNA sequence analysis according to Maxam and Gilbert[9].

Figure 3 shows the sequences determined around the splicing junctions. In the Lb gene the coding sequence is interrupted at the DNA sequences corresponding to the following amino acids in the amino acid sequence; IVS-1 at amino acid 32, IVS-2 between amino acids 68 and 69 and IVS-3 between amino acids 103 and 104. The sequences around the splicing junctions are consistent with the Breathnach–Chambon[10] rule because intervening sequences start with GT- and terminate with -AG. Furthermore, the sequences at the 5′ and 3′ ends of the intervening sequences agree with the consensus sequence for splicing junctions suggested by Lewin[11]. Thus plant genes share this feature with other eukaryotic genes[12].

Two intervening sequences have been found in all globin genes so far examined; in α globins the first interrupts between codons 31 and 32 and the second between codons 99 and 100. For β-globins the corresponding two intervening sequences are located between codons 30 and 31, and 104 and 105, respectively. In the Lb gene, interruptions occur at codons 32 and between 103 and 104.

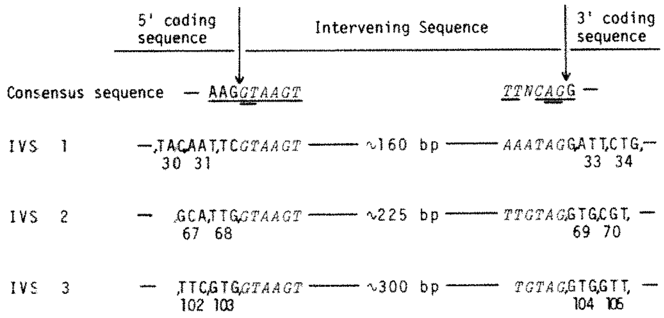

5' coding sequence
Intervening Sequence
3' coding sequence
Consensus sequence — AAGGTAAGT
TTNCAGG —
IVS 1 —,TAC,AAT,TCGTAAGT — ~160 bp — AAATAGG,ATT,CTG,—
30 31
33 34
IVS 2 — ,GCA,TTG,GTAAGT — ~225 bp — TTGTAG,GTG,CGT, —
67 68
69 70
IVS 3 — ,TTC,GTG,GTAAGT — ~300 bp — TGTAG,GTG,GTT, —
102 103
104 106

Fig. 3 Comparison of the position of Lb intervening sequence with that of a consensus sequence from Lewin[11]. Single underlines indicate a frequency of occurrence ≥45%, double underlines a frequency ≥95%. Numbers under codons refer to the number of the respective amino acid in the protein sequence. Codon 31 signifies asparagine, which identifies the isolated gene as a Lb c gene[1]. The DNA sequences were determined by the dideoxy chain-termination method described by Sanger *et al.*[8] and the Maxam–Gilbert technique[9].

By computer analysis Hunt *et al.*[13] have compared homologies of the amino acid sequences of globins including several Lbs. The observed similarity between the structures suggested that Lb and other globins have a common evolutionary origin, although they could be the result of convergent evolution. Inspection of the homology alignment of Hunt *et al.*[13] reveals that the coding sequences of all known globin genes, including the Lb gene, are interrupted by intervening sequences at identical positions. This finding supports the notion that all globin genes originate from a common ancestral gene.

The Lb gene is also interrupted between codons 68 and 69. This particular intervening sequence has not been found in the globin genes examined so far. Gilbert[14] has suggested that the presence of intervening sequences in DNA can speed evolution by allowing novel proteins to be constructed from the pieces of existing ones. Thus in the case of globin genes the central coding sequence codes for the parts of the globin chain which form the haem crevice. This particular correlation between coding sequence and function is not immediately apparent for the Lb gene, where the central coding sequence is split into two pieces of almost identical length. No known functions can be assigned to the coding sequences thus generated and consequently our result does not support the attractive hypothesis that a correlation always exists between a coding sequence and a particular function.

Being functional monomers, Lb and myoglobin are generally assumed to have evolved earlier from the common ancestral gene than α- and β-globins. Thus, the central intervening sequence either was present in the ancestral globin gene and subsequently lost during its evolution into α- and β-globins, or its insertion is a relatively new development in the evolution to Lb. Comparison of the Lb gene structure with that of myoglobin would help resolve this alternative but the DNA sequence of the latter remains unknown. In addition, several other Lb genes exist and their structures might in fact be different from the one determined here. Accordingly, we have cloned four other Lb genes and are investigating their detailed structures.

This work was supported by the Danish Research Council, NOVO Industry A/S and the Olga and Esper Boel Fond. We thank Drs B. Borkhardt and H. Delius for performing R-loop analysis, and Mrs Y. Gregersen and Mrs A. Marcker for technical assistance.

Received 2 February; accepted 9 April 1981.

1. Sievers, S. G., Huhtala, M.-L. & Ellfolk, N. *Acta chem. scand.* **B32**, 380–386 (1978).
2. Truelsen, E., Gausing, K., Jochimsen, B., Jørgensen, P. & Marcker, K. A. *Nucleic Acids Res.* **6**, 3061–3072 (1979).
3. Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).
4. Sullivan, D., Brisson, N., Goodchild, B., Verma, D. P. S. & Thomas, D. Y. *Nature* **289**, 516–518 (1981).
5. Leder, P., Tiemeier, D. & Enquist, L. *Science* **196**, 175–177 (1977).
6. Maniatis, T. *et al. Cell* **15**, 687–701 (1978).
7. Gronenborn, B. & Messing, J. *Nature* **272**, 375–377 (1978).
8. Sanger, F., Coulson, A. R., Bareil, B. G., Smith, A. J. H. & Roe, B. A. *J. molec. Biol.* **143**, 161–178 (1980).
9. Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560–564 (1977).
10. Breathnach, R., Benoist, C., O'Hara, K., Gannon, F. & Chambon, P. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4853–4857 (1978).
11. Lewin, B. *Cell* **22**, 324–326 (1980).
12. Sun, S. M., Slightom, J. L. & Hall, T. C. *Nature* **289**, 37–41 (1981).
13. Hunt, T. L., Hurst-Calderone, S. & Dayhoff, M. O. *Atlas of Protein Sequence and Structure* **5**, Suppl. 3, 229–251 (1978).
14. Gilbert, W. *Nature* **271**, 501 (1978).
13. Chow, L. T., Roberts, J. M., Lewis, J. B. & Broker, T. R. *Cell* **11**, 819–836 (1977).

A human proinsulin variant at arginine 65

D. C. Robbins*§, P. M. Blix*, A. H. Rubenstein*, Y. Kanazawa†, K. Kosaka† & H. S. Tager*‡

* Departments of Biochemistry and Medicine, The University of Chicago, Chicago, Illinois 60637, USA
†Third Department of Medicine, The University of Tokyo, Japan

Human diabetes can have several causes among which are the biosynthesis and secretion of mutant insulins with defective function. We recently described a patient whose diabetes resulted from secretion of an insulin in which leucine substituted for phenylalanine at position B24 (refs 1–4). Studies showing allelic variation in untranslated and intervening sequences of the human insulin gene (even in normal individuals) have documented a further and unexpected degree of variation in insulin gene structure[5,6]. We now describe the molecular abnormality of hyperproinsulinaemia in which large amounts of a proinsulin-like peptide are secreted[7–10]. The circulating proinsulin-like material from patients in one of the two known families with hyperproinsulinaemia[9,10] turns out to be a biosynthetic intermediate of proinsulin conversion in which the C peptide remains joined to the insulin A chain (see refs 11–13 for reviews of proinsulin biosynthesis and conversion). An amino acid alteration at position 65 (a position normally filled by arginine) apparently blocks conversion of the intermediate to biologically active insulin.

---

§Present address: Department of Medicine, University of Vermont, Burlington, Vermont 05401, USA.
‡To whom correspondence should be addressed.

0028-0836/81/260679—03$01.00

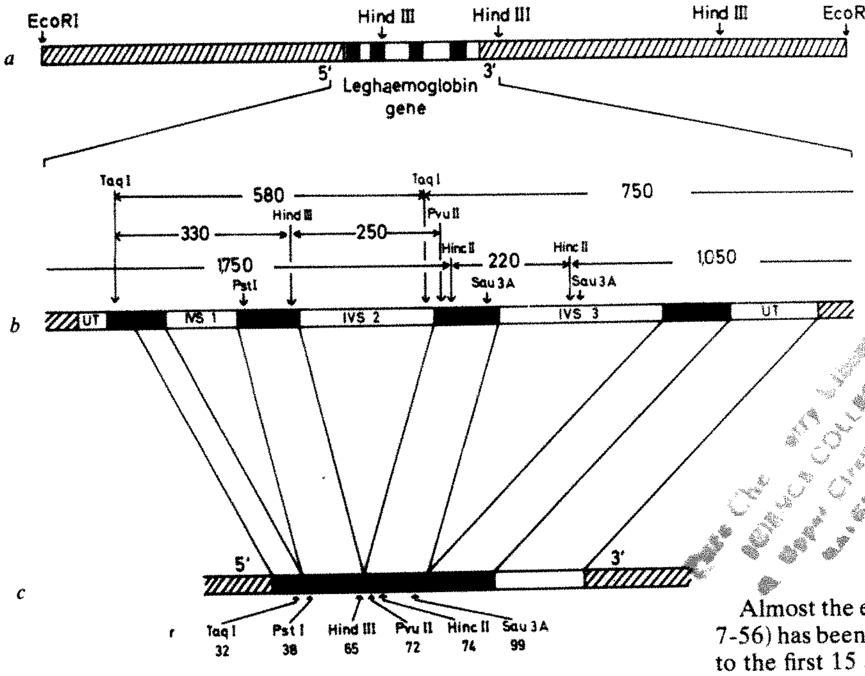


Fig. 1 Restriction cleavage map of the 7.5-kilobase soybean DNA fragment containing a Lb gene. Solid and open boxes represent coding and noncoding sequences, respectively; cross-hatched boxes represent non-Lb gene DNA sequences. UT 5′ and UT 3′ represent the sequence corresponding to the non-translated 5′ and 3′ ends of the Lb mRNA. IVS-1, IVS-2 and IVS-3 denote intervening sequences. Restriction endonuclease cleavage sites are indicated by arrows, and the distances between cleavage sites are given in base pairs. *a*, Map of the cloned *Eco*RI fragment of soybean DNA including the Lb gene. *b*, Expanded map of the Lb gene. *c*, Map of the Lb cDNA clone 7-56. Restriction endonuclease sites are here correlated with the number of the appropriate codon in the mRNA sequence. Cross-hatched boxes represent pBR322 DNA sequences.

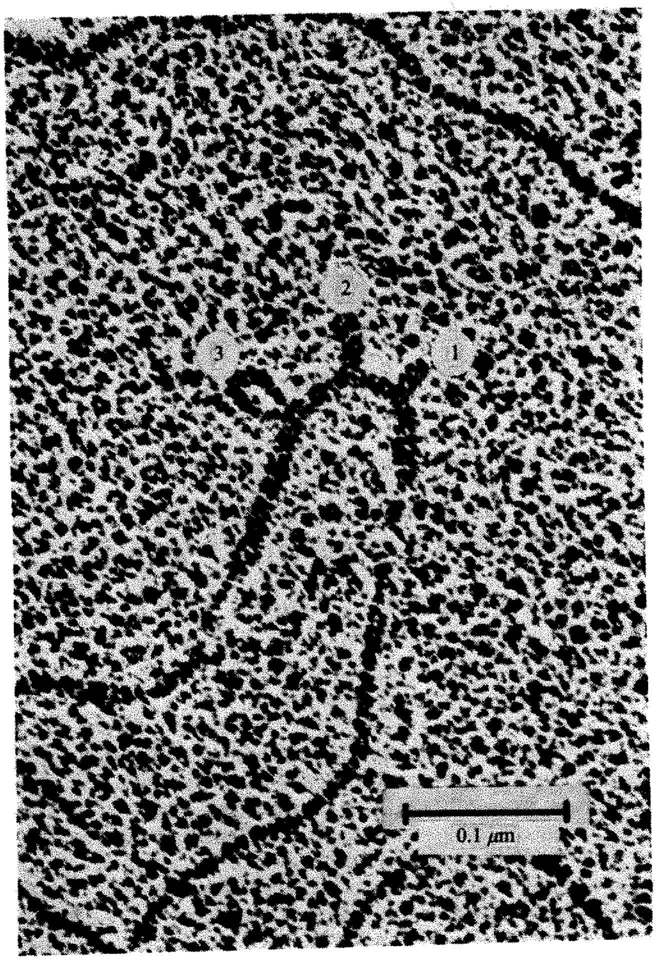


Fig. 2 Electron micrograph of the hybrid between Lb mRNA and the chromosomal Lb gene, formed by incubating the DNA and RNA in a buffer containing 80% formamide, 0.1 M HEPES *p*H 7.2, 0.4 M NaCl and 0.01 M EDTA for 2 h at 51 °C (ref. 15). The mixture was then diluted 10-fold into 30% formamide, 0.1 M Tris-HCl *p*H 8.5, 0.1 mg ml^{-1} cytochrome *c* and spread onto a water hypophase. After staining and shadowing with platinum, pictures were taken in a Philip EM 301. The arrows delineate the mRNA–DNA hybrid. 1, 2 and 3 refer to IVS-1, IVS-2 and IVS-3, respectively.

Almost the entire DNA sequence of a Lb cDNA clone (called 7-56) has been determined, lacking only that part corresponding to the first 15 amino acids of the Lb amino acid sequence. The DNA sequence is consistent with the amino acid sequence determined by Sievers *et al.*[1] for Lb c except that (1) Sievers *et al.*[1] determined the sequence -Phe-Val-Val-Lys- at amino acid positions 102–105 whereas we consistently find in this region a sequence corresponding to -Phe-Val-Val-Val-Lys-; and (2) Sievers *et al.*[1] found either Phe or Lys as a -COOH terminal whereas our equivalent base sequence corresponds to Ala.

Southern blotting analysis[3] of *Eco*RI digests of soybean DNA revealed the presence of at least seven different Lb genes in soybean, of length 1.4, 4.2, 5.5, 6.0, 7.5, 12 and 13 kilobases, respectively. Essentially similar results have recently been reported by Sullivan *et al.*[4].

To investigate the chromosomal Lb genes in more detail a library of soybean genes was constructed using λgt Wes·λB as cloning vector[5]. Soybean DNA was digested to completion with *Eco*RI and ligated to *Eco*RI/*Sst*-digested λgt Wes·λB. About 8 × 10^5 recombinant phage were screened[6]. Several plaques giving positive signals were purified and the DNA thus obtained digested with *Eco*RI and subjected to Southern analysis using nick-translated 7-56 DNA as a probe. In this way five separate clones carrying chromosomal Lb genes were detected. The sizes of the cloned soybean DNA fragments were 1.4, 4.2, 7.5, 12 and 13 kilobases, respectively. The clone containing the 7.5-kilobase fragment was subjected to further analysis by digestion with a variety of restriction endonucleases, from which a cleavage map for this particular Lb DNA gene could be constructed, as shown in Fig. 1. The sizing of the various restriction fragments suggested that the coding Lb DNA sequence was interrupted in three places with intervening sequences. To confirm this the 7.5-kilobase fragment was subcloned into pBR322, and the resulting clone allowed to form R-loops in the presence of Lb mRNA. The mixture was then analysed in the electron microscope with the result shown in Fig. 2; three loops are clearly indicated. The sizes of the three intervening sequences were estimated by direct measurement in the electron microscope and by sizing of the appropriate fragments obtained by digestion of Lb DNA with various restriction endonucleases. The results indicate that IVS-1 is ~160 base pairs long, IVS-2 230 base pairs and IVS-3 300 base pairs.

To establish the positions of the splicing points the 7.5-kilobase fragment was digested with *Sau*3A and the resulting fragments were subcloned into the M13 mp2 *Bam* vector developed by Messing[7]. Clones containing structural Lb DNA sequences were located by plaque hybridization using nick-translated Lb cDNA from clone 7-56. M13 clones containing Lb coding sequences were then subjected to DNA sequence analysis using the dideoxy chain-termination method developed by containing tryptic peptide of pp60^{src} was deduced: N-Leu-Ile-Glu-Asp-Asn-Glu-Tyr-Ala-Arg-C. The predicted amino acid sequence of this peptide in pp60^{src} includes potential cleavage sites for staphylococcal V8 protease, which cleaves at the carboxy-terminal side of glutamic acid residues[20]. We therefore digested pp60^{src} and ESV p80 with TPCK-trypsin, followed by V8 protease. In each case there was partial cleavage of the phosphotyrosine-containing tryptic peptide, generating a new peptide. Co-migration of the newly generated peptides (Fig. 1*h–j*) suggested that the cleavage products are identical in pp60^{src} and ESV P80. Similar results were obtained with P90 of Y73. Complete V8 protease digestion of the phosphotyrosine-containing tryptic peptide of ESV P80 yields a single new peptide in which the phosphotyrosine occurs at residue 1 (data not shown). Thus we suggest that a tryptic decapeptide containing the major tyrosine phosphorylation site of pp60^{src} is identical in the class III avian sarcoma virus polyproteins.

As the RSV *src* gene and the transformation-specific sequences of Y73 virus have no homology detectable by nucleic acid hybridization[2], and because the polyproteins of Y73 and ESV show no overlap with pp60^{src} in their methionine-containing tryptic peptides[3–5], it is assumed that the region of amino acid homology around the tyrosine phosphorylation sites represents a discrete and limited domain. We are now examining the extent of the sequence homology around these sites by cleavage with other proteases.

In addition to the homology in kinase target sites of pp60^{src}, P80 and P90 which we have described here, there is evidence that some features of these sites also occur in other transformation-specific proteins. We have found that the tyrosine phosphorylation sites on the class II avian sarcoma virus transformation-specific proteins, although not identical to classes I and III, are on tryptic peptides that are susceptible to digestion with V8 protease, yielding multiple partial cleavage products (Fig. 1*k*,*l*). Furthermore, polyoma virus middle-T antigen also has an associated protein kinase activity[21] which phosphorylates a tyrosine residue[22]. It is striking that the proposed sequence of middle-T contains a sequence of six glumatic acid residues followed by tyrosine; this tyrosine has been identified as the likely site of phosphorylation in the *in vitro* reaction (A. E. Smith, personal communication). These data and our results suggest that the occurrence of acidic amino acids adjacent to the phosphoacceptor tyrosine may be a general property of target sites for tyrosine-specific protein kinases, just as basic amino acid residues usually occur at serine phosphorylation sites[23]. These observations also support the hypothesis that such conserved sites may have a central role in regulating the function(s) of the viral transforming proteins as well as their normal cellular homologues.

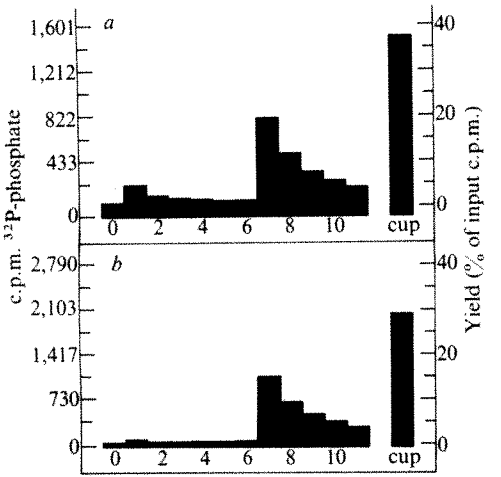


Fig. 2 Automic sequential Edman degradation of the ^{32}P-labelled tryptic peptides of polyproteins of Y73 and ESV avian sarcoma viruses. Tryptic peptides were first separated on a Sepheriso-b (C18) reverse-phase HPLC column. *a*, Peptide eluting at position 50–58 from Y73 P90 and *b*, peptide eluting at position 52–62 from ESV P80. ^{32}P-labelled peptides in 2 mg of apomyoglobin were dissolved in 500 µl of 18% formic acid. A 50-µl aliquot was taken for counting and the remainder applied to the spinning cup of a Beckman 890C sequencer with 3 mg of polybrene (Sigma). Sequencing was performed using a 0.1 M Quadrol buffer program. Conversion of the anilinothiazolinone derivatives, counting and determination of the carrier apomyoglobin residues (to ascertain that the sequencer had functioned properly) were as described by Smart *et al.*[27]. A zero cycle was included in each run by omitting the heptafluorobutyric acid cleavage step during the first programmed cycle. Recovery of the ^{32}P-labelled amino acid derivatives is incomplete on each cycle because of their strong affinity for polybrene.

During the course of this work we learned of similar observations on the phosphorylation sites of the Y73 and RSV proteins[24].

This work was supported by USPHS grants CA 13213 and CA 19725. J. N. holds a postdoctoral fellowship from the Leukemia Society of America. J. G. is supported by postdoctoral fellowships from the Fondation Rose et Jean Hoguet and NATO.

Received 27 January; accepted 9 April 1981.

1. Shibuya, M., Hanafusa, T., Hanafusa, H. & Stephenson, J. R. *Proc. natn. Acad. Sci. U.S.A.* **77,** 6536–6540 (1980).
2. Yoshida, M., Kawai, S. & Toyoshima, K. *Nature* **287,** 653–654 (1980).
3. Neil, J. C., De Lamarter, J. & Vogt, P. K. *Proc. natn. Acad. Sci. U.S.A.* **78,** 1906–1910 (1981).
4. Beemon, K. *Cell* **24,** 145–154 (1981).
5. Ghysdael, J., Neil, J. C. & Vogt, P. K. *Proc. natn. Acad. Sci. U.S.A.* **78,** 2611–2615 (1981).
6. Collett, M. S. & Erikson, R. L. *Proc. natn. Acad. Sci. U.S.A.* **75,** 2021–2024 (1978).
7. Levinson, A. D., Oppermann, H., Levintow, L., Varmus, H. E. & Bishop, J. M. *Cell* **15,** 561–572 (1978).
8. Hunter, T. & Sefton, B. M. *Proc. natn. Acad. Sci. U.S.A.* **77,** 1311–1315 (1980).
9. Kawai, S. *et al. Proc. natn. Acad. Sci. U.S.A.* **77,** 6199–6203 (1980).
10. Feldman, R., Hanafusa, T. & Hanafusa, H. *Cell* **22,** 757–765 (1980).
11. Neil, J. C., Ghysdael, J. & Vogt, P. K. *Virology* **109,** 223–228 (1981).
12. Ghysdael, J., Neil, J. C., Wallbank, A. M. & Vogt, P. K. *Virology* **111,** 386–400 (1981).
13. Erikson, R. L., Collett, M. S., Erikson, E. & Purchio, A. F. *Proc. natn. Acad. Sci. U.S.A.* **12,** 6200–6204 (1979).
14. Sefton, B. M., Hunter, T. & Beemon, K. *J. Virol.* **33,** 220–229 (1980).
15. Oppermann, H., Levinson, A. D. & Varmus, H. E. *Virology* **108,** 47–70 (1980).
16. Erikson, R. L., Collett, M. S., Erikson, E., Purchio, A. F. & Brugge, J. S. *Cold Spring Harb. Symp. quant. Biol.* **44,** 907–917 (1979).
17. Levinson, A. D., Oppermann, H., Varmus, H. E. & Bishop, J. M. *J. biol. Chem.* **255,** 11973–11980 (1980).
18. Smart, J. E., Oppermann, H., Czernilofsky, A. P. & Bishop, J. M. *Proc. natn. Acad. Sci. U.S.A.* (submitted).
19. Czernilofsky, A. P. *et al. Nature* **287,** 198–203 (1980).
20. Houmard, J. & Drapeau, G. R. *Proc. natn. Acad. Sci. U.S.A.* **69,** 3506–3510 (1972).
21. Smith, A. E., Smith, R. & Paucha, E. *Cell* **18,** 915–924 (1979).
22. Eckhart, W., Hutchinson, M. A. & Hunter, T. *Cell* **18,** 925–933 (1979).
23. Krebs, E. G. & Beavo, J. A. *A. Rev. Biochem.* **487,** 923–959 (1979).
24. Patschinsky, T. & Sefton, B. M. (in preparation).
25. Neil, J. C., Breitman, M. & Vogt, P. K. *Virology* **108,** 98–110 (1981).
26. Smart, J. E. & Ito, Y. *Cell* **15,** 1427–1437 (1979).
27. Smart, J. E., Lewis, J. B., Mathews, M. B., Harter, M. L. & Anderson, C. W. *Virology* (in the press).

The structure of a chromosomal leghaemoglobin gene from soybean

Erik Østergaard Jensen, K. Paludan,
J. J. Hyldig-Nielsen, P. Jørgensen & K. A. Marcker

Department of Molecular Biology and Plant Physiology, University of Aarhus, DK-8000 Århus C, Denmark

Leghaemoglobin (Lb) has only been found in the nitrogen-fixing root nodules of legumes symbiotically associated with *Rhizobia*. Soybean nodules contain three major forms of Lb called a, c_1 and c_2, respectively. The amino acid sequences of the three Lbs have been determined recently[1]. Lb a differs from the two c varieties in six amino acids, while Lb c_1 and c_2 differ in only one amino acid. We have recently reported the isolation of a hybrid plasmid containing soybean Lb cDNA sequences[2]. Using one such clone as a probe, we have isolated five different soybean DNA fragments containing chromosomal Lb genes from a soybean DNA library which was constructed using lambda λgt Wes·λB as vector. The detailed analysis of one of the isolated chromosomal Lb genes revealed the presence of three intervening sequences. Alignment by homology of all known globin amino sequences shows that the splicing points of the first and the third intervening sequences in the Lb-coding sequence are located in the same positions as the two interruptions found in all other known globin-coding sequences.

0028-0836/81/260677—03$01.00

BOOK REVIEWS

Reason and unreason in assessing risk

Eric Ashby

Society, Technology and Risk Assessment. Edited by J. Conrad. Pp.304. ISBN 0-12-186450-2. (Academic: 1980.) £15.20, $35. *Dealing with Risk: The Planning, Management and Acceptability of Technological Risk.* Edited by R.F. Griffiths. Pp.144. ISBN 0-7190-0819-0/ 0-470-27136-1. (Manchester University Press/Halsted: 1981.) £11.50, $19.95.

LAST November the Royal Society held a discussion meeting on the assessment and perception of risk. Among the speakers were engineers, toxicologists, psychologists, mathematicians, biologists, administrators; but there was no representative of the profession which makes its living out of risk-assessment: the actuaries who calculate insurance premiums. If the experts in this profession are not cautious enough, the firm loses money. If they are too cautious, the firm loses business. The impression left (upon me, at any rate) by this discussion was that research on risk-assessment, though much of it is interesting and sophisticated, is not yet a reliable tool for people who have to make decisions in the public interest about hazardous projects. Statisticians can give accurate estimates of familiar risks. Social psychologists can give plausible descriptions of the way these risks are perceived. What we need is more evidence from the people who may lose their reputations or their jobs if they misjudge risks. With this in mind I welcomed the invitation to read two more books on this subject.

The book edited by Conrad is a record of a workshop organized by the Battelle Institute. There are 20 papers by authors from different countries. Part I covers familiar ground; it deals with the methods used in risk-analysis. Nothing fresh here, except a perceptive criticism of the "revealed preference theory" popularized by Starr in a paper in *Science* 12 years ago (**165**, 1232–1238). Parts II, III and IV cover the social and political consequences of risk-assessments. Here there is a good deal of interesting and controversial opinion.

Risk-assessment is only one of several operations necessary for the much broader discipline of decision-analysis, and it is a pity that it has been split off and isolated from the parent subject. For when you have made your statistical assessment (of, for instance, lives that would be saved by the wearing of seat belts in cars), and your guess as to how motorists perceive this risk, you still lack data essential for making a decision. The politician wants an answer to the question: "What is an acceptable risk?". But there is an antecedent question to that one: "Who, in a pluralistic democracy, has the right to decide what is an acceptable risk?". When the issue arises in Parliament, this is the sort of question which MPs ask. And what is the purpose behind a statement about risk-assessment? Is it to impose safety standards for which there is a public demand? Is is to reassure the public that some decision already taken, for reasons nothing to do with safety, need not cause disquiet? Then there is the assumption, one I have myself made, that public perceptions of risk are irrational (for example the person who willingly indulges in some practice which exposes him to a 10^{-3} risk of death but who campaigns fanatically against some other practice which would expose him to a 10^{-7} risk of death). Irrational, no doubt, on the criteria applied professionally by scientists, but most of the British public are not scientists and their irrationalities are as significant for the decision-analysis as are the hard data from the statisticians. To exclude their views from the risk-benefit balance sheet would be to cook the books. Social intuition may appear to be a woolly concept compared with fault-tree-analysis, but it is likely to carry more weight in Parliament and in Whitehall.

Owing to considerations such as these, a good deal of research on risk-assessment seems peripheral to the issues that have to be decided by developers or planning authorities. This is the opinion of some of the contributors to the Battelle Institute workshop. What emerges from these papers is the impression that (in the words of Volker Ronge, from Germany)

> . . . decision-making processes have apparently come out to be risky in that they are no longer capable of producing sufficient social consent to outcomes that have to a large amount been "decided upon" outside the political realm.

The book edited by Conrad, then, enlarges our awareness (if not our understanding) of the sociological and political ingredients in decision-analysis. The book edited by Griffiths is a more pragmatic set of papers, most of them written by people who really have to make decisions about risky projects and who work at the coal-face of public opinion. It is therefore refreshingly practical. A safety adviser from industry sets out clearly how he has to compose his advice. No use to pretend that there is zero risk. Sooner or later some accident will happen and there will be fatalities. There is a temptation to rely on compared risks ("we should not spend money on raising the height of the handrails if the risk of falling over them is smaller than the risk of being harmed by toxic chemicals in the atmosphere"). But of course this sort of argument is not acceptable to the industry or to the public. The decision-maker is driven back to crude cost-benefit analysis: the cost of saving a statistical life. Thus (an example from Conrad's book) in the Netherlands it was considered whether to introduce earth leakage switches into domestic wiring systems. It would reduce the number of electrocutions by a factor of ten; but since the present risk is only 5×10^{-7} per year, the cost per life saved would be of the order of $2 millions.

It is no wonder that this sort of arithmetic is not commonly disclosed to the public in planning enquiries, nor, if it were, would it make much impression. In two chapters on hazard and risk in planning, two planners (one from the Cheshire County Council and the other from the Halton Borough Council in Cheshire) explain exactly what procedures they follow when there are hazards in the project being planned. The book is worth reading for these two contributions alone; for the experience of these two practical men may do something to remedy the most serious problem in risk-assessment, namely the alienation of the public toward those who have to decide, on the public's behalf, what is an acceptable risk. Risk, as Stephen Cotgrove points out in the concluding essay in *Dealing with Risk*, "is not just a statistical calculation. It is also a moral judgement about defensible conduct". It is the legitimacy of decisions made about risk, more than their accuracy, which is being called in question by public interest groups and which causes exasperating delays in what are often desirable developments. The open, empirical and unsophisticated procedures described by the planners from Cheshire, coupled with their willingness to take the public into their confidence, may do more to restore the legitimacy of planners than would volumes of research on risk-assessment.

These two books, together with the book *Societal Risk Assessment* (reviewed in *Nature* **289**, 827; 1981) should dispel any illusion that the problem: How safe is safe enough? is within sight of solution. But why, oh why, have we not yet heard evidence from actuaries who calculate insurance?

Lord Ashby is Chancellor of Queen's University, Belfast, and a Fellow of Clare College, Cambridge.

Ninth stage in the long march of Joseph Needham

R.G.W. Anderson

Science & Civilisation in China, Vol.V:4. By Joseph Needham. Pp.772. ISBN 0-521-08573-X. (Cambridge University Press: 1980.) £48, $105.

Dr Needham's great compilation at this stage occupies a substantial length of bookshelf — already more, for example, than Thorndike's monumental *A History of Magical and Experimental Science.* Though this volume is the ninth to appear, a recently issued "State of the Project" pamphlet states that the completed series will comprise 20 books. There will be three further volumes on chemistry and chemical technology after the currently reviewed volume, the third to appear.

Three major areas are covered by this book: chemical techniques, alchemical theory and comparative macrobiotics. The first is largely involved with a description of distillation processes. Here, as with many other topics treated by Needham in the series, the reader is not simply provided with an elucidation of the Chinese tradition, but with the best comparative account available for a number of cultures. However one main argument has now been shown to be inaccurate: the suggested greater efficiency of the Eastern stills to the simple Western types (retort or alembic). Needham proposes that the relatively late date for the distillation of alcoholic spirits in the West (13th century AD) can be explained by the ineffectiveness of the condensation surface of the retort and alembic. In China, he maintains, the more efficient stills account for evidence of spiritous liquors in the 7th century AD. This argument has been demolished in a recent paper in *Ambix* (**27**, 69; 1980) (of which Needham, it must be admitted, is co-author) in which stylized Western retorts, Mongolian stills and Chinese stills were tested experimentally and were shown to be equally efficient at alcohol distillation. How odd, then, that there is no positive evidence of earlier distillation of spirits in the West as there is of perfume oils and petroleum fractions.

Much less is said about other processes and the associated apparatus, though there is an interesting discussion of reaction-chambers (which do not have an exact Western counterpart). Cunningly designed water-cooled examples are illustrated, redrawn from the *Chin Hua Chhung Pi Tan Ching Pi Chih* ("Confidential Instructions on the Manual of the Heaven-piercing Golden Flower Elixir") of 1225 AD. So bizarre are some of these that one wonders if they were really constructed.

The second section, "Alchemical Theory", discusses the application of Taoist philosophy to an understanding of chemical materials and their reactions. To the Western mind of today, the explanations of the Chinese alchemists bear little or no relationship to current scientific understanding; Needham himself says that "the dominant goal of proto-scientific alchemy was contemplative" (p.244) — though this must be reconciled with a later statement that theoretical alchemy was "a deductive proto-science on quite the same level as medicine, acoustics or magnetic geomancy" (p.298).

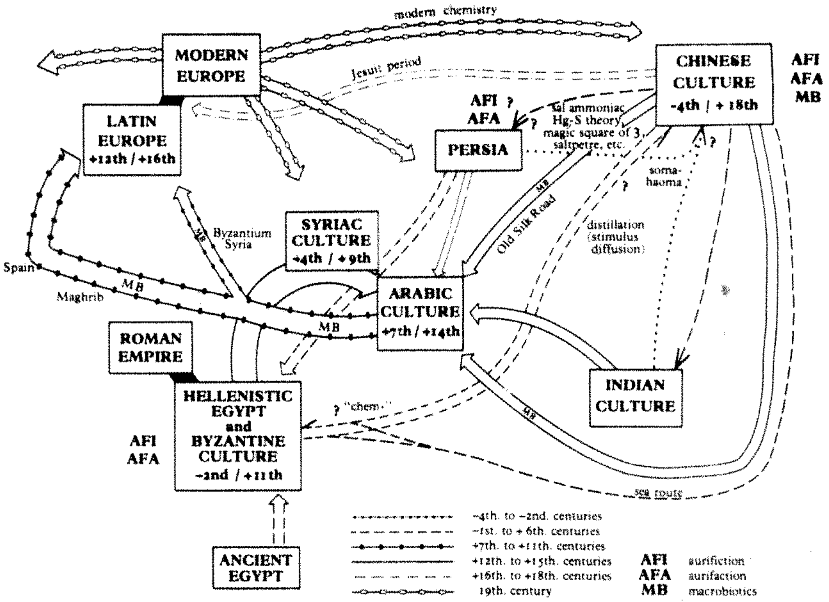


Fig. 1538 of *Science & Civilisation in China*, Vol. V:4.

Chart illustrating the multifocal origins of proto-chemistry, to be seen as superimposed roughly upon a map of the Old World. See p.504 of the book for further details.

In the final part of the book, macrobiotics is defined as longevity or material immortality attained with the aid of chemical knowledge. Here a comparison is made of Chinese alchemy, Hellenistic proto-chemistry (macrobiotics played little part in Graeco-Egyptian culture) and Arabic and Latin alchemy. Most importantly, the interactions and hence influences are probed. Thus it is shown that the early chemistry of the Arabs was not exclusively Graeco-Egyptian in origin, but included contributions from China, India and Persia. A summing-up is provided in the form of a chart (reproduced above), illustrating the complex multifocal origins of proto-chemistry.

The reader of this authoritative and rewarding book is left with the impression of two themes pursued relentlessly by Needham. First, routes of influence are constantly sought, the chances of independent discovery of the same solution of the same problem in geographically distant places only being considered when all else fails. This can produce some rather absurd, tongue-in-cheek suggestions such as the relationship between the Mongolian still and the Irish poteen still. Second, is the effort made to claim primacy for the Chinese on every occasion. Thus in a passage attempting to pre-date Western spirit distillation by the Chinese, Needham follows some rather ambiguous early evidence with the revealing remark "We still have plenty of time before us ere we reach the + 12th century, time of the first distillation in the West".

Yet it is the highly personal approach which helps to make this prodigious work so enjoyable. How delightful and revealing, for example, is the seemingly careless comment: "Resting one day in 1968 at a wayside auberge in the South of France, I read again the text of the *Tabula Smaragdina* and felt so much the Chinese flavour of it that I started translating it to see how it would look in that language". □

R.G.W. Anderson is Keeper of the Department of Chemistry at the Science Museum, London.

Bright future of NSE

B. E. F. Fender

Neutron Spin Echo. Lecture Notes in Physics, 128. Edited by F. Mezei. Pp.253. ISBN 3-540-10004-0. (Springer-Verlag: 1980.) $19.50, DM 33.

Neutron spin echo (NSE) techniques represent the single most important innovation in neutron scattering methods in recent years. The idea originated with F. Mezei and has been skilfully realized at the Institut Laue–Langevin in a full working instrument by J. Hayter and Mezei.

Conventional neutron inelastic scattering instruments for studying the dynamics of materials are conceptually simple; neutrons of a particular energy are selected by a monochromating crystal or a velocity selector and, after interaction with a sample, the neutron energy is redetermined. The instrument design is dominated by the relatively low flux of neutron beams and this problem is particularly acute when a high-energy resolution is required, for example to probe slow motions in polymer chains. The breakthrough which the spin echo technique represents is that the velocity of the neutron is determined from the Larmor precession of the neutron spin in a magnetic field. In this way the resolution of the energy change of scattering does not depend on the monochromatization of ingoing or outgoing neutron beams, and very small energy changes down to 10^{-8} eV can be measured.

This book, which is based on a workshop at the ILL, is in three parts. The first deals with the method from the point of view of both theory and its practical realization. The discussion is sufficiently detailed for the specialist to appreciate as a review, while at the same time it provides a useful introduction for the intending practitioner.

The non-specialist will probably be more interested in the second part which gives a range of examples. Even though most of the measurements are preliminary, this section well illustrates how improvements in neutron instrumentation generate fresh scientific opportunities. The chapters which are likely to be the best indicators of future research are those on the dynamics of micelle solutions, polymer solutions, spin glass dynamics and superfluid ^{4}He, although the reader must beware of the rapidity of developments in these and other fields. The third section deals primarily with future instrumentation; it will serve for some time as food for thought for the experiments of tomorrow. Apart from concluding comments by the editor, there are also interesting chapters on the use of spin echo methods to discriminate against thermal diffuse scattering and on the application of NSE techniques in order to measure phonon line widths.

In sum, then, this is a thoroughly worthwhile contribution to a good series.

B. E. F. Fender is Directeur Adjoint at the Institut Max von Laue-Paul Langevin, Grenoble.

Reluctant leaf-eaters of Barro Colorado

R.D. Martin

The Foraging Strategy of Howler Monkeys: A Study in Primate Economics. By Katharine Milton. Pp.165. ISBN 0-231-04850-5. (Columbia University Press: 1980.) $20, £13.

It is generally agreed that C.R. Carpenter's 1934 monograph (*Comp. Psychol. Monogr.* **10**, 1-168) on mantled howler monkeys, *Alouatta palliata*, marked a significant transition in primate field studies, from essentially anecdotal accounts to serious scientific investigations. His study was conducted on the small island of Barro Colorado (BCI), an area of about 1,500 hectares of rainforest isolated during the construction of the Panama Canal in 1914. The Smithsonian Tropical Research Institute has, over a period of several decades, sponsored numerous fundamental studies of the fauna and flora. In the half century since Carpenter's publication, quite a few studies have been conducted on the primate species inhabiting BCI, and the howler monkey population (now numbering over 1,000 individuals) has been re-studied and re-censused at fairly regular intervals. *Alouatta palliata* therefore counts among the best-known primate species with respect to natural behaviour and ecology.

The first systematic primate field studies were, inevitably, largely concerned with basic data collection and primarily qualitative. As work has progressed, there has been increasing emphasis on quantification of both behavioural and ecological parameters, and most modern studies are specifically designed to answer well-defined questions. There has also been increasing integration of laboratory studies with field work as more specific issues, such as nutritional and toxic properties of natural foodstuffs, have come into focus. All of these modern trends are admirably exemplified in Katharine Milton's *The Foraging Strategy of Howler Monkeys*, which provides a fine example of the great advances made since Carpenter's pioneering work. Appropriately, Milton's main 14-month field study was also conducted on BCI, on the direct descendants of the monkeys studied by Carpenter.

The howler is one of the largest of the New World monkeys (Ceboidea) and has been widely regarded as the most folivorous. Milton's quantitative, year-round data confirm that, on average, the mantled howler spends more time eating leaves than either fruits or flowers. Yet the howler lacks any obvious specialization of its digestive tract to extract particular benefit from an abundant gut flora in processing the leaves in its diet (viz. breakdown of cell wall material; detoxification). For instance, there is no sacculation of the stomach as is found in the specialized leaf-eating monkeys of the Old World (Colobinae), and the caecum is not notably enlarged. Using carefully-collected quantitative data on both behaviour and ecology, Milton sets out to define the "foraging strategy" of the howlers and to explain the lack of obvious digestive tract specialization. In a clear and economically written account, she gradually constructs a picture of the mantled howler as a species which neatly balances considerable selectivity in feeding against tight budgeting of energy spent in travel. It emerges that the BCI howlers preferentially feed on relatively clumped fruit and flower sources when available and complement their diet by selectively feeding almost exclusively on young leaves (which, as a rule, contain more protein and less cell-wall material than mature leaves). Seasonal items are particularly important and the howler modifies its feeding pattern over the annual cycle, feeding most on leaves (including a small proportion of mature leaves) when other items are scarce. Despite its concentration on food items with a spatially and temporally patchy distribution, *Alouatta* is a relatively sluggish primate, resting for about 65% of daylight hours. It seems to save energy by travelling directly to a few "pivotal" trees (usually fruit-sources) each day, picking up the rest of its food with little extra travel. A wide range of different plant species (at least 73 per troop) is exploited and the overall effect is to limit travel distance to the low level of 443m per day. In line with many other primate species, the howler is most active, and has the most diversified diet, during the season of greatest availability of preferred foods. Surprisingly, although the vegetation on BCI is dense tropical rainforest, there is a marked seasonal pattern (consistent from year to year). Still more surprisingly, it is the "dry" season on BCI that is apparently most favourable in terms of availability of preferred foods of the howlers.

A brief review cannot do justice to the theoretical scope and importance of this monograph. In addition to the central theme of foraging strategy, the text broaches a number of important areas — such as the relation between social organization and feeding ecology, the co-evolution of plants and animals and the evolution of plant secondary compounds as defence systems — which are at the spearhead of current primate field studies. This is an outstanding book, with only a few minor shortcomings (for example inadequate discussion of problems of bias in data collection and insufficient integration of the results with information from other quantitative studies of howlers and other primates), and it deserves to be widely read by zoologists, primatologists and anthropologists. As a companion volume to Pierre Charles-Dominique's *Ecology and Behaviour of Nocturnal Prosimians* (1977), Milton's monograph sets a very high standard indeed for any further books on this topic from Columbia University Press. □

R.D. Martin is Reader in Physical Anthropology at University College and Visiting Professor in Zoology at Birkbeck College, University of London.

Is there nothing new in primatology?

Robert A. Hinde

The Social Life of Monkeys and Apes, 2nd Edn. By S. Zuckerman. Pp.511. ISBN 0-7100-0681-8. (Routledge & Kegan Paul: 1981.) £17.50, $55.

IN 1932, Zuckerman's *Social Life of Monkeys and Apes* was far ahead of its time. It contained scholarly reviews of animal sociology and of the physiology of the oestrous and menstrual cycles. It was remarkable in attempting to integrate physiology with behaviour. Several chapters reviewed the author's own pioneering observations of baboons in South Africa and Hamadryas baboons in the London Zoo. From these he drew conclusions about monkey societies, and about the nature of social and sexual responses, which stimulated much research. Re-reading it now one marvels at the scope of the data synthesized, at the keen critical sense of the author and the number of issues, now the subject of detailed research, which appear there in embryonic form. To mention just one, the importance of "contact comfort" for non-human primates, re-discovered by Harlow in 1959, was discussed by Zuckerman in 1932.

The book stands as a landmark because it helped usher in a new way of studying monkeys. Of course we now know much more about both endocrinology and non-human primates, and new theoretical perspectives have developed. This unrevised re-issue of Zuckerman's book (the addition of a postscript and seven appendices containing reprinted essays or reviews hardly justifies the label of "Second Edition") provides an opportunity to assess the progress that has been made. Three generalizations may serve as examples. First, Zuckerman held that the nucleus of all primate societies is "the family party, consisting of an overlord and his harem" (p.314). Whilst true of hamadryas, recent field studies have shown that some species are monogamous (e.g. gibbons, marmosets), and that superficially polygynous ones show a great diversity of structural types. Whilst in some species there are close relationships between particular males and particular females, in others male–female liaisons are much more temporary (e.g. baboons, macaques).

Second, in 1932 Zuckerman regarded sex as the main force that holds groups together (e.g. pp.59, 314). But whilst the primary bonds in some polygynous species are between male(s) and females (gorilla), more usually they are between mothers and their daughters. In the latter case, each troop may consist of one or more matrilines (a matriarch, her daughters and their daughters) to which males become temporarily attached (macaques, baboons), or are quite peripheral (patas). Zuckerman's emphasis on sexual bonds as paramount seems to stem from a view that single sex groups do not have bonds (pp.58–59), perhaps based on a presupposition that the "family", containing male(s) and females, is necessarily central to any society. The third issue is Zuckerman's emphasis on male dominance as the major principle determining social structure. Dominance is of course important, but we now know that females can be dominant to males for at least some of the time (talapoin, squirrel monkey), and that social structure depends also on other issues, including a correlate of genetic relatedness. Indeed field primatology is now in an exciting stage, with new theoretical insights matched by field data revealing new subtleties in social structure. Primatology, for long lagging behind other branches of biology, is now generating important general principles.

Journals supplement

On October 1 *Nature* will publish a review supplement devoted to new journals which have appeared between January 1978 and May 1980. Publishers and societies are invited to submit sample copies for consideration for review; for details see *Nature* **291**, 599.

That the book now needs amendment is no reflection on the first edition: we see further now in part because we stand on the shoulders of giants, and for a related reason must also acknowledge that some of our own conclusions may later need extension or correction. But the unfortunate postscript to this re-issue brings no such feeling of humility about the progress of science. It conveys instead the impression of an attempt to discredit not only whole areas of research developed since 1932 but also the contributions of young workers within those fields, and to show that primatology has made little progress in 50 years. In doing so, Zuckerman goes to rather remarkable lengths. He attacks quantification and in particular the field workers' practice of stating the extent of their data base (p.363) and use of statistical techniques (pp.370–371), apparently in part because he mistakenly perceives monkey behaviour to be almost infinitely variable (pp.371, 375–377, 403). He ignores most experimental work, and descends to describing most field workers as "novices with little experience of scientific research or method", incidentally distorting a semi-popular review of a sketchy article on the problems they face (p.366). Some of his criticisms are based on citations which inaccurately represent the original source (e.g. p.371, lines 33–39). He misunderstands functional questions, rejecting research into ultimate causes as ending "either in empty statement or in tautology" (p.404) and, although earlier regretting the lack of ecological data on primates (p.31), he labels most subsequent research as "tortuous and intellectually unrewarding" (p.375). His emphasis on the variability of monkey behaviour leads him almost to deny inter-species differences (pp.375, 379): he ignores most modern evidence on the diversity of social systems, and tries to discredit the rest. He seems unaware of the advances consequent upon the kin selection hypothesis.

Now of course not all primatology (nor indeed all endocrinology) has been of the highest quality, and apparently promising theoretical alleys can prove unrewarding. Some criticisms, especially those directed at the popular writing of Morris, Ardrey and sometimes Lorenz, are justified. But Zuckerman's postscript gives the impression of being almost wholly destructive, and contrasts with the original book in not attempting a synthesis. This is a pity: the re-issue of this once influential book might have been accompanied by a constructive appraisal so that it, like the original, facilitated future progress. ☐

Robert A. Hinde is a Royal Society Research Professor and Honorary Director of the Medical Research Council Unit on the Development and Integration of Behaviour, Madingley, Cambridge.

Star of the Galaxy

D. Lynden-Bell

Oort and the Universe: A Sketch of Oort's Research and Person. Edited by Hugo van Woerden, Willem N. Brouw and Henk C. van de Hulst. Pp.210. ISBN hbk 90-277-1180-1; ISBN pbk 90-277-1209-3. (Reidel: 1980.) Hbk $29, Dfl.55; pbk $12.95, Dfl.25.

OORT was the heir of a great Dutch tradition in astronomy but under his guidance it has flowered to an extent unforeseen by his predecessors. In this book we have a sketch of Oort's major contributions to galactic and cometary astronomy and catch a glimpse of a man whose foresight and quiet insistence gave Holland the world's most powerful radio telescope at Westerbork.

The book is an eightieth birthday tribute to Oort from 25 leading astronomers who have worked with him or have been close colleagues. Some concentrate on Oort's scientific influence, some on the political difficulties that Oort overcame to get the money for a major instrument and some on the science that it opened to exploration.

However, this is not a *Festschrift* in which leading scientists contribute samples of their work in honour of the recipient; rather, it is a survey of history in the making, concentrating on Oort's period and his interests. Here we have some basic material from which a historian might begin his study of the way the character of a great scientist allows him to further science through his influence as well as his own research. Anyone who has studied galactic astronomy, radio astronomy or cometary astronomy will find here some fascinating pieces of history. B. Lindblad's picture of galactic rotation was that the stars with greater velocity dispersion had slower rotation about the galaxy because of their greater pressure support. Each such subsystem was originally considered to rotate uniformly. It was Oort who saw that stars of very low velocity dispersion must rotate differentially and who then went on to show that this motion could already be seen in the observed proper motions.

Later, it was Oort who saw the importance of finding spectral lines in the radio region and sent van der Hulst on the search that led to his great discovery of the 21 cm line. Until reading this book I had always regarded it as a sad irony of fate that the Dutch, with that prediction but with no wartime experience of radar, just missed being the first to observe the line. Much credit goes to Muller for bringing the techniques up to date so quickly, but the gap was too great and, as he points out here, his task was made quicker by learning of the techniques used in the US which led Ewen to the first detection. However, Oort knew what should be done with the line to probe the astronomy of the Galaxy and so it was the Dutch who won the next round. It is sad that G. W. Rougoor who collaborated with Oort in some striking papers died when so young, but I am sure he would too have been a contributor.

However, if I may choose a favourite among Oort's discoveries, it is his determination of the total density of visible and invisible matter in our region of the Galaxy through his work on the gravity field. As Blaauw describes, this work opened the door to the astronomy of the invisible which has grown to be a major subject over the past ten years and in which the Westerbork telescope has played a major part. Some of the highlights of the discoveries are described in the well-illustrated contribution by Allen and Ekers, while van der Laan describes Oort's part in the use of Westerbork observations in his new role as Professor Emeritus.

I give the last word to van Bueren who contributes a well-written and perceptive article on Oort's role in Dutch astronomy:

> The kind and courteous Oort holds fast to his opinions with unbelievable firmness. He therefore finally always wins his argument, with colleagues as well as with research-financing organisations and government officials.

Perhaps this has been Oort's secret of success, which has helped him to put Holland in the forefront of modern astronomy.

D. Lynden-Bell is Professor of Astrophysics at the Institute of Astronomy, Cambridge.

Rheology from Russia

J.J. Hermans

Rheology of Polymers: Viscoelasticity and Flow of Polymers. By G.V. Vinogradov and A. Ya. Malkin. Pp.467. ISBN 3-540-09778-3. (Springer-Verlag: 1981.) $57.90, DM 98.

THIS translation of a revised Russian text published in 1977 contains an admirable treatment of the rheology of polymers. It begins with a comprehensive introduction to the mathematics of deformation and flow, which covers both linear and nonlinear constitutive equations.

More than 30 pages on linear viscoelasticity are followed by rather brief, but pertinent, coverage of nonlinear theories. There are excellent sections on the newtonian and the non-newtonian viscosity of polymers in relation to molecular weight, polydispersity, branching, concentration and entanglements. A remarkably extensive discussion of about 40 pages covers normal stresses. There are further chapters on birefringence, rubber-like behaviour in flow (i.e. large recoverable deformations), the effect of fillers and the rheological properties of polymer blends.

The authors have wisely decided to emphasize concentrated solutions and melts, these being the systems of greatest practical interest. The well-known molecular models and theories that have been developed for very dilute polymer solutions are mentioned, but play no great role in the discussion of either the theory or the experimental data. In other words, the treatment is mostly phenomenological. Many empirical equations are discussed but, when at all possible, these are related to the phenomenological theory.

Particularly praiseworthy is the balance between theory and experiment. This is perhaps best illustrated by the fact that Chapters 2–7 contain about 180 figures showing experimental results.

Inevitably, in a book which covers such a vast field, one will notice omissions. For example, whereas empirical relations between non-newtonian viscosity and shear stress are mentioned, there is nothing on the requirement of zero slope at the limit of zero shear stress. To mention another example, in the chapter on uniaxial extension the investigator will find that many literature references are missing; conversely, it can be said that the value of the book is enhanced by some references with which the English, and even more so the North American reader is often unfamiliar.

The book will be a good introduction for anyone wishing to learn about this field. At the same time it will be a valuable reference work to scientists already experienced in the subject.

J.J. Hermans is Emeritus Professor of Physical Chemistry at the University of North Carolina.

SCIENTIFIC BOOKSHOP

H.K. LEWIS can supply works in all branches of Pure and Applied Science. Catalogues on request. Please state interests.

SCIENTIFIC LENDING LIBRARY

Annual Subscription from £7.50.
(Available in U.K. only)

Reduced rates for multiple subscriptions.

Prospectus post free on request.

Quarterly List of New Books and new editions added to the Library sent post free to subscribers regularly.

H.K. LEWIS ...
& Co. Ltd.
136 GOWER STREET,
LONDON, WC1E 6BS

Telephone: 01-387 4282
Telegrams: "Publicavit,
London, WC1E 6BS."

Circle No.14 on Reader Enquiry Card.

Bring Your Library Up-To-Date With This ANS Edition...

J.M. Cleveland
MONOGRAPH

The Chemistry Of Plutonium

Approx. 680 pages 6" × 9"
Hardbound $49.00

American Nuclear Society
555 North Kensington Avenue
La Grange Park, IL 60525 USA

Please take my order for ______ book(s).
☐ Payment in full enclosed.
☐ Bill me and I accept postage and handling charges.

Name ______________________
Organization ______________________
Street ______________________
City ______________________
State ______________ Zip __________

Circle No.02 on Reader Enquiry Card.

Social scientists and sociobiologists get their lines crossed

Jon Seger

Sociobiology Examined. Edited by Ashley Montagu*. Pp.272. ISBN hbk 0-19-502711-6; ISBN pbk 0-19-502712-4. (Oxford University Press: 1980.) Hbk £12.50, $19.95; pbk £2.95, $5.95.

Sociobiology Examined is a collection of 16 articles, half of which have been or will be published elsewhere. The authors are listed below. According to the dust jacket, "the articles . . . examine the claims [E.O.] Wilson makes for [human] sociobiology [in *Sociobiology* and *On Human Nature*] and, by a critical examination of the evidence (or lack of it) on which these claims are based, assess their validity". In only a few of the articles is there a sustained attempt to do this. Some are mostly polemical, in places irresponsibly so; some are mostly about subjects other than sociobiology. For example, Layzer's article turns out to be a detailed critique of Piaget's theory of cognitive development. Sheehan's is a review of recent books on political philosophy by Bernard Henri-Levy and Alain de Benoist, neither of whom appears to have the slightest idea what sociobiology is or says, except that someone told them it supports their opinions. Most of the articles stick more closely to the stated subject than do these, but taken as a whole they give the impression that "sociobiology" is more of a Rorschach test than it is a theory or even a well-defined point of view. Midgley comes to a similar conclusion in her essay on the first few years of the "debate" on sociobiology. Although the articles present a bewildering range of views, two major themes come up fairly often. One concerns the supposed social, political or ethical implications of sociobiological theories, and the other the issue of genetic determinism.

In agreement with Plato, Hobbes, Rousseau and Wilson, most commentators on sociobiology seem to believe that a theoretically and empirically adequate description of human nature necessarily implies a morally compelling prescription for the organization of human society. The lone dissenter in *Sociobiology Examined* is Mackintosh, whose article was first published as a review of *On Human Nature* (*Science* **204**, 735; 1979). "The gap between 'is' and 'ought' remains as wide today as it was when Hume pointed it out 200 years ago". Mackintosh quotes Wilson's claim that "a correct application of evolutionary theory . . . favors diversity in the gene pool as a cardinal value", and notes that this is only one of several places where Wilson states or implies that a value can be derived from a fact. But as Mackintosh also notes, Wilson's "radical" critics appear to be even more committed to the naturalistic fallacy than he is. It is truly incredible that in six years of heated "debate" there has been hardly any serious discussion of this problem.

The question of genetic determinism is much more complex. As Midgley points out, "determinism" is simply a belief in the efficacy of causes. Thus unless you are a dedicated supernaturalist, or believe in at most one cause, it is silly to accuse someone of being a "determinist". As a matter of logic this is true enough, but the social scientists appear to have something more concrete in mind when they speak of genetic or biological determinism. Unfortunately, few of the authors in *Sociobiology Examined* ever state explicitly what they mean by phrases like "genetic basis", but by the end of the book it has become clear that many of them believe sociobiology attempts to map differences of behaviour, whether between individuals or societies, onto corresponding differences of genotype. In other words, any variability that has a "biological" explanation is necessarily heritable. This belief can be seen clearly in the article by Simon.

> Behavior that can plausibly be designated as innate or instinctive must be stable throughout a range of environments . . . . Human social behavior . . . varies considerably from one social group to another. Social science is typically interested in explaining these variations . . . . If the determinants of the specific differences among a widely disparate range of behavior are not biological but sociocultural, it will not explain any of them to subsume them all under a common biological rubric.

Sociobiology does, in part, attempt to explain behavioural variation. But does it really argue for the *heritability* of such variation? Those who think that sociobiological theory derives from traditional behaviour genetics say "yes"; this group includes most of the authors in the present collection, but very few sociobiologists. Those who identify sociobiology with the newer work on sex ratios, sexual dimorphism, kin-directed altruism, life-history strategies and the like, say "no"; this group includes most sociobiologists, but in the present collection only Barkow, who notes that "Evolutionary biology does not require . . . a direct link . . . between a particular gene and a particular behavior, even if careless reading (and writing) of sociobiology may occasionally permit that impression". There is no contradiction in this as long as different levels of causation are properly distinguished. Barkow's discussion of the levels of causation needed to explain human behaviour is well worth reading. He observes, as have many others, that selection may often favour behavioural flexibility of a kind "mediated by learning preferences". There would be little point to flexibility unless it enabled an individual to recognize and seize unique opportunities. Consistent with this general expectation, animals in learning experiments tend to master some associations much more easily than others. If sociobiology is really about the evolution of learning and other environmentally mediated developmental processes, then at the level of differences between individuals it is, ironically, a theory of *environmental* determination. Individual differences map onto differences of environment, not of genotype, although they do so according to a logic that is "genetic" in the sense that it reflects innate, evolved characteristics of the species. Thus only at the level of differences between species is sociobiology inherently a theory of genetic determination. Given these contrasting views of "sociobiology", and given the distinctions on which they turn, any simple reference to a genetic substrate, basis, cause, determination or the like, unless carefully specified, is almost certain to be literally meaningless. Because these key terms are seldom specified in the pages of *Sociobiology Examined* the reader often cannot tell just where a given piece of criticism is supposed to strike.

Without a doubt the sociobiologists are partly responsible for this widespread confusion regarding levels of causation and explanation. Mackintosh rightly criticizes Wilson for shifting between different senses of "the notion of genetic determination", and more generally for not giving us "a serious and sustained attempt to analyze in just what sense our behavior or social organization is genetically determined . . .". But if Wilson's writing is careless or confused about something so fundamental, why should that of his sharpest critics be even more so? Surely critics have a responsibility to expose weaknesses, rather than to hide them.

Several observers of the sociobiology controversy have remarked that the participants seem to be talking past each other. Unfortunately, *Sociobiology Examined* does more to document this situation than to change it. The book will be of interest to historians and sociologists of science. But for the most part it will be of little help to those who are still wondering how the study of animal behaviour may ultimately improve the study of human behaviour, and vice versa. □

Jon Seger is at the Biological Laboratories, Harvard University.

---

* Articles by Jerome H. Barkow, S.A. Barnett, Derek Freeman, Stephen Jay Gould, Marvin Harris, James C. King, David Layzer, N.J. Mackintosh, Mary Midgley, Ashley Montagu, Karl Peter and Nicholas Petryszak, Stephen Rose, Thomas Sheehan, Michael A. Simon and S.L. Washburn.